University of Alberta

Creating a Developmentally Appropriate Self-Report Instrument of Children's Optimal Challenge During Physical Activity

by

James Lloyd Mandigo ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Faculty of Physical Education and Recreation

Edmonton, Alberta

Fall, 2001

0-612-68967-0

University of Alberta

Library Release Form

Name of Author: James Lloyd Mandigo

Title of Thesis: Creating a Developmentally Appropriate Self-Report Instrument of Children's Optimal Challenge During Physical Activity

Degree: Doctor of Philosophy

Year this Degree Granted: 2001

#33-130 St. Augustine Dr.
St. Catharines, ON
L2P 3X2
Canada

Date: June 20th, 2001

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled: Creating a Developmentally Appropriate Self-Report Instrument of Children's Optimal Challenge During Physical Activity submitted by James Lloyd Mandigo in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dr. Linda Thompson

Dr. Marcel Bouffard

Dr. Brian Nielsen

Dr. Michael Enzle

Dr. Todd Rogers

Dr. John Dunn

Dr. Stuart Biddle

May 29th, 2001

Dedication

For Pop

"We think of you and smile"

Abstract

Creating developmentally appropriate items is critical when conducting self-report research with children. The common practice of taking adult questionnaires and simply modifying the wording can often result in unnecessary sources of measurement error (Brustad, 1998). The purpose of the studies outlined in this dissertation was to create developmentally appropriate items for a self-report instrument of optimal challenge for use with children and to present various sources of construct validity evidence. A total of 27 (15M; 12F) children in study one were interviewed about their physical activity experiences. Items representing three subscales (skill equals challenge; skill > challenge; and, challenge > skill) were developed from participants' responses. The relevance of these items was then reviewed by an international group of experts ($n = 9$) in study two. Following revisions, the similarities and differences between items were then rated by 15 (3M; 12F) children. Multidimensional scaling techniques were used to determine which items were most similar and which items were least similar. For the final study, the items from study three were used to create the Children's Perceptions of Optimal Challenge Instrument (CPOCI). A total of 95 (57M; 38F) children used the CPOCI to rate the degree to which they felt they were optimally challenged during an aiming task in a field setting. A series of multivariate and univariate analyses were used to examine the validity and reliability of the CPOCI. Based upon a unified perspective of validity (Messick, 1989), construct validation evidence of the instrument was supported through various types of validity. Recommendations for future validation studies and the potential pedagogical and research implications of the CPOCI are discussed.

Acknowledgements

A very special thanks to Dr. Linda Thompson who was not only great advisor, but is a tremendous role model and person. You have set a high standard and have challenged our field to grow and to be proud of itself. You have made a difference not only in my life, but also in thousands of other students and teachers across the country.

Thanks as well to my supervisory committee: Dr. Marcel Bouffard, Dr. Brian Nielsen, and Dr. Mike Enzle for challenging me to look outside the box and to ask tough questions. Your feedback and encouragement along the way was greatly appreciated.

Thanks to those who served as external examiners (Dr. Todd Rogers, Dr. John Dunn, Dr. Gordon Walker & Dr. Stuart Biddle) who took time out of their busy schedules to provide me with very valuable feedback.

Thanks also to Nicholas Holt and Jamie Covey for all your assistance with data collection, analysis, and support. This would have been a very difficult road to travel without your continual support and friendship.

Thanks goes out to all the participants in this study for their honesty, time, and commitment to the various studies in this dissertation.

Thanks to my "new" colleagues at Brock University who continually encouraged me to keep going and who offered advice and support along the way.

Finally, I'd like to thank Nate for her support, patience, and love. You gave me the support and confidence to keep going. You also taught me that there was more to life than just this project and academia and I look forward to spending the rest of my life with those most important to me: my family.

TABLE OF CONTENTS

PAGE

PAGE

PAGE

PAGE

LIST OF TABLES

TABLE	DESCRIPTION	PAGE

LIST OF FIGURES

CHAPTER 1

Introduction

On January 13, 1999, Michael Jordan announced his retirement from professional basketball stating that he had run out of personal challenges. As a result of his inward desire to seek out and conquer new and more difficult challenges within the game of basketball, he achieved feats that earned him the title of being "super-human" on the court. This desire for activities that challenge our abilities is not limited to elite athletes. According to three organismic theories of motivation (i.e., Csikszentmihalyi, 1975; Deci & Ryan, 1985; Harter, 1978a), humans have an intrinsic desire to seek out and participate in activities that foster their self-development. Such activities are said to be optimally challenging.

Individuals are optimally challenged when they perceive the challenge(s) of the activity to be balanced with their abilities to do the task(s) (Csikszentmihalyi, 1990a; Reeve, 1996; Weiss, 1986). Humanistic theories of motivation hypothesise that when participants feel optimally challenged during an activity, they are more likely to have a quality subjective experience and be intrinsically motivated to take part in the activity at that time and in the future due to enhanced perceptions of competence and enjoyment (Reeve, 1996). However, if exposed to continuous imbalances (i.e., skill ≠ challenge), participants may become frustrated or bored which may eventually lead to their withdrawal from the activity (Csikszentmihalyi, 1975; 1990a). The importance of providing children with positive and enjoyable physical activity experiences such that they will want to remain physically active has been stressed by many as one of the most important considerations for keeping children active (Scanlan & Simons, 1992; Wankel, 1993). Providing settings and experiences for optimally challenging activities can be one of the ways to stimulate and maintain this intrinsic motivation for children.

In the late 1950's and early 1960's, the behaviourists' assumptions about motivation were challenged by researchers such as White (1959) and Maslow (as cited in Miller, 1993). These leaders in motivational research adopted an organismic perspective which suggested that humans freely interact and consequently bring about change within their environment (Deci, 1975). Based on this approach, humans have an internal desire to experience positive affect through enhanced competence and self-actualisation. From these organismic and humanistic theories of motivation developed the importance of optimal challenge. Researchers such as Harter (1978a), Deci (1975), and Csikszentmihalyi (1975) have all identified the importance of optimal challenge within their theories of intrinsic motivation. As well, school physical education programs currently stress the importance of matching the challenges of an activity with a child's developmental abilities to maximise success and skill development (Buschner, 1996; Graham, 1992; Morris & Stiehl, 1999; National Association for Sport and Physical Education, 2000; Weiss & Bressan, 1985). However, despite the theoretical underpinnings that have been developed within the area of motivation and the importance of understanding optimal challenge in structuring developmentally appropriate physical activity environments, limited empirical research currently exists to demonstrate the antecedents of fostering optimal challenge and the psychological and behavioural outcomes of such an experience. The overall purpose of this dissertation is to initiate a research program that will lend itself to a better understanding of the experience of optimal challenge in children's physical activity environments. For the remainder of this chapter, an extensive summary and critique of the current research and theoretical perspectives on optimal challenge is presented. Based on this review, recommendations for creating a developmentally appropriate self-report instrument to tap into the optimal challenge construct are provided. Using this review and the recommendations as the foundation for this dissertation, the results from four separate studies are presented in paper format (Chapters Two to Five). More specifically, these studies are intended to provide preliminary construct validation evidence for an instrument capable of providing information on the extent to which children are optimally challenged during physical activity participation.

Review of Literature

The concept of optimal challenge is embedded within three intrinsic motivation theories: Theory of Optimal Experience (Csikszentmihalyi, 1975; 1990a); Cognitive Evaluation Theory (Deci, 1975; Deci & Ryan, 1985); and, Competence Motivation

Theory (Harter, 1978a). According to each of these theories, being optimally challenged is imperative to facilitating intrinsic motivation.

Theory of Optimal Experience

Originally, Csikszentmihalyi (1975) was interested in understanding what makes an activity enjoyable. His original seminal work with talented artists examined why individuals would devote hours towards their paintings without seemingly caring about the extrinsic rewards (e.g., financial incentives). He went on to investigate what motivated individuals such as amateur athletes, chess masters, rock climbers, dancers, and composers of music to pursue their activities when the external rewards were not forthcoming. What he found was that it was their sense of intrinsic enjoyment that motivated them to participate and to continue their participation.

According to Csikszentmihalyi (1990a), individuals who reach a flow state have a highly motivating experience. This state is characterised by a sense of control, a merging of action and awareness, a loss of self-consciousness, a transformation of time, informative feedback, and sufficient skill to take on a challenge. As this suggests, one of the prerequisites to being in flow is that an individual must be optimally challenged. The flow state is highly motivating and fosters a desire to seek out similar experiences in order to fulfil the needs of the self to grow and develop (Csikszentmihalyi, 1988a). In its purest form, "the purpose of flow is to keep on flowing" (Csikszentmihalyi, 1975, p. 47). When individuals experiences a flow state, it is believed to have a positive impact on the quality of life (Singer, 1996). As people increase their skill level, they are motivated to seek out increasingly challenging activities. Individuals are likely to become bored with the activity if it becomes too easy. Conversely, if an activity that is perceived as being far too challenging for their current skill level, they will experience anxiety and frustration.

Initially, Csikszentmihalyi (1975) argued that certain individuals and/or activities may be more likely to experience or produce flow experiences respectively. Individuals who hold an autotelic personality are said to be more likely to experience flow than others. Autotelic refers to self-rewarding goals and comes from two Greek words: "auto" meaning self and "telos" meaning goal (Csikszentmihalyi, 1990a). Autotelic personalities are personal characteristics that are self-directing or self-rewarding and enable an individual to sustain and enjoy an activity with the end result being the attainment of a flow state (Csikszentmihalyi, Rathunde, & Whalen, 1993). Such personalities are characterised by individuals who are able to enjoy what they are doing regardless of external rewards. Hypotheses put forth by some have suggested that dispositional characteristics such an individual's goal orientation, perceived competence, game goals, concentration, state and trait anxiety, sense of autonomy, task preference, and self-efficacy all impact upon the development of flow states (Jackson & Roberts, 1992; Kimiecik & Stein, 1992; Kowal & Fortier, 1999; Rea, 2000; Stein, Kimiecik, Daniels, & Jackson, 1995). Although an argument has been made that some individuals are born with an autotelic disposition, it can also be argued that people can also learn how to balance the challenge of the activity with their skill level (Csikszentmihalyi, 1988a).

The environment in which individuals participate also has a major impact on whether participants will experience a flow state. Autotelic activities contain patterns of action within the activity that maximise immediate intrinsic rewards to the participant (Privette. 1983) and meet the conditions previously described for a flow state to occur. Although a flow state typically occurs in structured activities in which the level of challenges and skills can be varied and controlled (Csikszentmihalyi, 1988a), the experience of flow can occur in almost any activity environment provided certain conditions are met. These include individuals setting achievable goals, evaluating these goals over an extended period of time, being able to focus on the task, providing opportunities to develop skills necessary to interact with the episode, and creating challenges that are easily modified to avoid boredom and anxiety (Csikszentmihalyi, 1990b).

Within a sport environment, Kimiecik and Stein (1992) recommended that environmental and situational factors should be considered when trying to facilitate a flow state. For example, self-paced sports such as golf that involve closed skills are more conducive to flow experiences than open skill environments where the performance is dependent upon someone else (e.g., team-mate passing the ball, opponent returning a tennis shot) or variable conditions. McInnman and Grove (1991) suggested that sports which offer fast and direct feedback such as golf or baseball enable athletes to see quickly and clearly whether they have reached their goal of hitting the ball in the middle of the fairway or have laid down the sacrifice bunt.

Athletes' perceptions of the sport experience may also play an important role in fostering flow state experiences. For example, conditions where athletes perceive they have choice, clarity, challenge, and commitment are more likely to lead to a flow state than conditions where the players feel that they are being controlled, do not know their role, are placed in situations in which their skill is far advanced or not yet developed to meet the challenges posed by the activity and do not feel part of the "team" (Kimiecik & Stein, 1992). For children, Kleiber (1981) suggested several strategies to help facilitate flow experiences in organised sport. These strategies centre around providing a non-threatening, self-determined environments that enables participants to enhance their own personal development through optimally challenging activities that they choose to do.

In summary, Csikszentmihalyi's theory of optimal experience provides a solid theoretical framework to understand the experience of optimal challenge and its psychological outcomes. His operational definition of flow (i.e., skill = challenge) was described by Deci and Ryan (1985) as the best measure of optimal challenge currently available. The weakness with this operational definition of flow, however, is that it does not match the way the theory describes a flow state. Even Csikszentmihalyi (1992) recognised this weakness when he cautioned that one can not just measure flow on a scale by saying it is only the balance between above average skill and challenge. This balance is simply a required condition for reaching a flow state. The balance between skill and challenge is better operationalised as optimal challenge and is one of the antecedents for reaching a flow state. When a person is optimally challenged, it does not guarantee that she/he will experience flow; however, a person can not experience flow without this balance of high challenge and skill. Also, the theory of optimal experience has suggested that there may be certain dispositional and situational variables that help to predict whether a person will feel optimally challenged and experience a flow state. Knowing these conditions are useful for researchers when designing a model where potential antecedents are presented to help predict the occurrence of an optimally challenging experience.

Cognitive Evaluation Theory

According to Deci and Ryan's (1985) self-determination theory, individuals have an intrinsic desire to feel autonomous, competent and related. Such needs motivate individuals to seek out optimally challenging events in which competence can be enhanced in self-determined (i.e., free choice) environments. For example, when the challenge of an activity is too high or too low relative to their ability levels, individuals seek out situations that are optimally challenging (Deci, 1975). Such motives can be undermined, however, when individuals perceive that they are controlled by external factors such as grades or evaluation (Deci & Flaste, 1995). Deci and Ryan (1985) presented cognitive evaluation theory (CET) as a sub-theory of self-determination theory as a way to help explain factors in specific social contexts that serve to undermine or enhance a participant's intrinsic motivation (Ryan & Deci, 2000). Traditionally, intrinsic motivation has been operationalised behaviourally by the amount of time that participants freely spend at a task or subjectively by the use of questionnaires (e.g., McAuley, Duncan & Tammen, 1989) that measure the amount of interest and enjoyment participants have for an activity (Deci & Olson, 1989). Using self-determination theory as the foundation, CET contains four propositions that help to explain which variables enhance and which variables undermine intrinsic motivation.

According to Proposition 1, intrinsically motivating activities are autonomous or self-determined (Frederick & Ryan, 1995). When individuals participate in an activity in which they feel as if they had some choice and control of the process to reach personal goals, intrinsic motivation will be enhanced. Conversely, when individuals participate in an activity and feel externally controlled, intrinsic motivation is likely to be decreased. Goudas, Biddle, Fox, and Underwood (1995) tested this hypothesis with the use of different teaching styles in a physical education class. Girls (n= 24; M = 13 years of age) reported significantly higher amounts of intrinsic motivation when their track and field instructor offered students a number of choices throughout the lesson (i.e., differentiated teaching style) as opposed to conditions where all the decisions were made by the instructor (i.e., direct teaching style). Other studies have also demonstrated that when participants are given more control and choice within the activity environment, intrinsic motivation is likely to be enhanced (Deci & Olson, 1989; Deci & Ryan, 1985).

Proposition 2 states that intrinsic motivation is enhanced by feelings of competence and optimal challenge. Deci (1975) first outlined the importance of optimal challenge on intrinsic motivation by suggesting that because humans are motivated to be self-determined and competent, they seek out optimally challenging situations. When individuals take part in activities that are neither too hard nor too easy relative to their skill level and they feel self-determined in the process, their competence is enhanced. This enhanced competence in turn leads to individuals feeling motivated to participate. Despite the empirical links that have been demonstrated between perceived competence and intrinsic motivation (e.g., McAuley, Wraith, & Duncan, 1991; Vallerand, 1983; Vallerand & Reid, 1984; Whitehead & Corbin, 1991), there exists a lack of empirical evidence within the CET framework to demonstrate the impact of optimal challenge on perceived competence and self-determination. Anshel, Weinberg, and Jackson (1992) did conduct a study with undergraduate students in which they found that participants reported higher levels of intrinsic motivation after they had learned a more complex task (i.e., went from juggling 2 bean bags to 3 bean bags). However, they did not examine this relationship with respect to perceived competence.

The third proposition alludes to the functional significance of extrinsic factors that can be viewed along a continuum as to their impact on intrinsic motivation. The effect of these extrinsic factors can also be viewed from a social perspective regarding the impact that others have on an individual's intrinsic motivation (Deci & Ryan, 1994). Extrinsic factors that are perceived as being informational with respect to one's perceived competence (i.e., provides effectance feedback in the context of choice) promote intrinsic motivation whereas extrinsic factors which are perceived as being controlling (i.e., are used to control behaviour) and/or amotivational (i.e., are perceived as conveying incompetence and helplessness) undermine intrinsic motivation. The most controlling and amotivational extrinsic factors are classified as external regulation. These factors are perceived as controlling an individual's behaviour by demands or contingencies external to the person. For example, parents who pay their children for scoring goals in ice hockey or for other performance-related criterion often undermine their children's intrinsic motivation for participation in the sport. Introjected regulations are perceived as controlling behaviour by demands or contingencies inside the person. For example, children who go to swimming practice because they will feel guilty if they do not go are controlled by feelings that they "should" do something or because they would feel ashamed if they did not. Identified regulations are behaviours chosen because the person identifies with the importance of the activity. Children and adolescents who state that they participate in sports to improve fitness are likely influenced by identified regulations because they value the outcome associated with being physically active. Integrated regulations are very similar to intrinsic motivation because they are experienced as "wholly free" due to the regulation being integrated with the person's sense of self. However, unlike intrinsic motivation which is characterised by behaviours performed because the person is interested in the behaviour itself, integrated behaviour is performed freely because the outcome is meaningful and important (Deci & Ryan, 1994).

Research relevant to the sport and physical activity field concurs with the theory that extrinsic rewards that are perceived as being controlling and/or amotivational undermine intrinsic motivation. For example, Orlick and Mosher (1978) randomly assigned children from a day camp to either a reward or no reward condition. Participants were then observed for the amount of time they spent on a stabilometer. Four days later, participants were called back and those who had received no reward spent significantly more time on the stabilometer during the second trial than on the initial trial. Those who received a reward, however, spent significantly less time on the stabilometer during the second trial compared to the initial trial. Ryan, Mims, and Koestner (1983) reported similar findings with undergraduate students who received a reward doing puzzles. They found that students who were promised a reward for their participation in the study spent significantly less time working on the puzzles during "free time" than students who did not receive a reward. However, the rewards given with the intent of controlling behaviour (i.e., will receive monetary reward should you perform to our standards) undermined intrinsic motivation significantly more than rewards given with the intention of informing participants of their abilities (i.e., will receive monetary reward if you do well on the puzzles).

The issue of perception is another important concept. The same reward may be viewed by some as being controlling and for others as being informational. For example, E. Ryan (as cited in Frederick & Ryan, 1995) interviewed intercollegiate athletes who held athletic scholarships. For football players who normally received them, the scholarships reduced their intrinsic motivation to play football because they were viewed as being controlling. For wrestlers and for female athletes, the scholarships increased intrinsic motivation because they were viewed as being informational regarding their ability because at that time, such scholarships were rare for both groups of athletes.

The final proposition of CET refers to intrapsychic regulation. Here, a person's perception and orientation towards the activity influences their intrinsic motivation. People who hold a task-orientation will likely be more intrinsically motivated because they take part in an activity for the enjoyment of the task while those who hold an ego-orientation will likely not be intrinsically motivated because they are influenced by a pressure to appease self-esteem. For example, people who feel they have to score a goal to set a new record likely hold an ego-orientation towards scoring goals as opposed to those who play the game to the best of their abilities because they enjoy it. These orientations are often influenced by experience. If children, for example, are continually being told that their self-worth is contingent upon performance in an activity, they will soon adopt an ego-orientation. Conversely, if children are allowed to participate in freely chosen activities for the sake of participating in the activity, a task-orientation is more likely to emerge.

Duda, Chi, Newton, Walling, and Catley (1995) provided empirical evidence to support this theoretical connection. They found task-oriented goals to be significantly correlated with intrinsic motivation and ego-oriented goals to be negatively correlated with intrinsic motivation. Likewise, Goudas et al. (1995) found that those with a high task- and low ego-orientation were more likely to report high levels of intrinsic motivation than others with different goal orientations (i.e., high ego, low task). Within a competitive situation, scenarios that foster an ego-orientation tend to undermine intrinsic motivation. For example, children assigned to a competitive group (i.e., beat the scores of other children) for participation on a stabilometer had significantly lower levels of intrinsic motivation (i.e.. time on task) than those in an intrinsic mastery (i.e., do the best you can) condition (Vallerand, Gauvin, & Halliwell, 1986).

In summary, CET provides a sound humanistic perspective of motivation and offers insight into the significance of optimal challenge. Although the construct of optimal challenge is embedded within the second proposition of CET, limited empirical evidence exists to support its role within CET. Future studies need to be conducted to examine the relationship optimal challenge has with other variables embedded within CET such as self-determination, perceived competence, sense of relatedness and belonging, extrinsic rewards, and intrapsychic regulation. Perhaps one of the limitations to studying the impact of optimal challenge is the lack of an operational definition. Neither of the traditional questionnaires used in CET studies of sport (i.e. the intrinsic motivation inventory or the situational motivation scale) have a measure of optimal challenge. Therefore, despite suggestions of potential ways to create optimally challenging states and psychological outcomes of such a state, there currently does not exist an empirical measure for optimal challenge within CET.

Competence Motivation Theory

In what has now become a classic paper on motivation from a cognitive perspective, White (1959) suggested that all motivation is not influenced strictly by biological drives within the organism. Humans have a need to seek out activities that enhance their sense of competence and which are enjoyable at the same time. For this, White (1959) coined the term "effectance motivation". Although White's work drastically changed the way we have come to understand and research human motivation, he failed to provide a way in which to operationalise and measure intrinsic motivation. It was here that Susan Harter advanced White's work and provided developmentally appropriate instruments (e.g. Harter, 1981b; 1985; Harter & Pike, 1984) to test the many components of what is now called Competence Motivation Theory (CMT).

According to Harter (1981a). individuals seek out intrinsically motivating and optimally challenging activities in which they can demonstrate their level of competence.

The greater effectance motivation, the more likely people will seek out optimally challenging activities (Reeve, 1996). When children take part in activities that challenge them at a level that is balanced with their skill level, experience success, and receive informative and positive feedback from adults, they are more likely to experience an increase in their level of perceived competence (Weiss, 1986). However, this may be easier for some than others. Some children hold what Harter (1981b) refers to as an intrinsic motivational orientation. These individuals are more likely to experience increases in perceived competence because of their desire to challenge themselves, to be genuinely interested in the task for their own personal reasons, and to feel self-determined in their pursuit to improve their abilities.

Within the domain of sport and physical activity psychology for children, considerable efforts have been to better understand the role that perceived competence plays on adherence and on how to foster enhanced perceived competence in participants. The majority of the research has demonstrated that those who take part in sports and physical activities tend to have higher levels of perceived competence than those who drop-out or who do not participate (Biddle, 1997). This may be due to children with high perceived competence choosing to participate in order to demonstrate their competence, or that those with low perceived competence do not have the same opportunities to succeed within a sport environment. Whatever the connection, it is widely accepted that sport and physical activity programs for children need to offer every child, regardless of skill level, an opportunity to feel competent. This need or desire to enhance or demonstrate competence is due to the high positive affect that is associated with high perceived competence. For example, those who have high levels of perceived competence tend to report higher levels of self-esteem (Weiss, 1986), a desire to continue participation (Gibbons & Bushakra, 1989; Klint & Weiss, 1987), self-determination (Harter, 1978b), fun/enjoyment (Ommundsen & Vaglum, 1991; Weiss, Bredemeier, & Shewchuk, 1985), self-efficacy (Harter, 1978b), and effort (McKiddie & Maynard, 1997) than those with low levels of perceived competence.

In summary, CMT provides the third theoretical link to understanding optimal challenge. Harter's (1978a) extension of White's (1959) concept of effectance motivation resulted in the development of a social-cognitive theory of motivation which has addressed the importance of providing children with optimally challenging experiences to enhance their perceived competence.

Findings Pertaining to Optimal Challenge Research

Although all three of the intrinsic motivation theories discussed so far address the importance of optimal challenge, only a handful of studies have empirically examined psychological antecedents and/or outcomes of optimal challenge or have directly examined how children describe and experience the construct. The first known empirical evidence to point out the importance of optimal challenge was conducted by Harter (1974) within the framework of CMT. Harter (1974) asked 40 fifth and sixth grade children to solve anagrams at various difficulty levels (3, 4 and 5 letters). She found that the greatest level of enjoyment was derived from solving more challenging anagrams as opposed to easily solved problems. In follow-up interviews after the experimental phase, 67.5% of the participants said they would rather repeat challenging or unsolved anagrams than the ones they were able to solve easily.

In a subsequent study, Harter (1978b) asked children to solve anagrams again, but this time, introduced an evaluative condition (i.e., performance was graded versus performance was not graded) and a six letter anagram. Like the previous study, enjoyment was positively correlated with anagram length up to the five-letter solution. However, a curvilinear relationship between task difficulty and enjoyment was found when the amount of smiling decreased for the six letter anagram. Examining the impact of being evaluated, participants in the "no grade" condition chose more difficult anagrams and had higher smiling scores than those in the "grades" condition. In follow-up interviews, 80% of students in the "no grades" condition said they would have chosen the easier anagram had they been evaluated while 75% of students in the "grades" condition indicated they would have chosen harder anagrams had they not been evaluated on their performance. Similar findings have also demonstrated that the combination of grades imposed by an impersonal evaluator, a focus on the correct solution (product-orientation), and the salience of social comparisons undermine children's interest and enjoyment of the learning process and influence their preference for challenging activities (Harter, 1992).

Although Harter's initial research was important to suggest a curvilinear relationship between enjoyment and task difficulty and the impact that situational variables such as rewards (e.g., grades) have on children's choice of difficulty level, it did not take into consideration the initial ability level of the subjects. Danner and Lonky (1981) addressed this issue by dividing 90 children in kindergarten and Grades 1, 2 and 4 into three groups based on their cognitive skill levels. They then set up three stations that were intended to improve upon the children's cognitive skills. Children were given the choice of how long they spent at each station. The results showed that children reported higher levels of interest and spent more time at stations that were either equal to or slightly higher than their skill level. An inverted-U relationship was also demonstrated between interest ratings and perceived difficulty level. Tasks rated as either too easy or too hard were not rated as interesting as tasks that were matched according to the children's skill level. The effects of rewards (i.e., a good work certificate) were also investigated. The results demonstrated that when intrinsic motivation, as measured by time-on-task and interest level, for an optimally challenging activity was initially high, rewards decreased the children's motivation. When intrinsic motivation for a task perceived as being more difficult was initially low, rewards had little effect on motivation. Combined, the results from this study supported the importance of providing children with choice and respecting their decisions when creating an optimally challenging environment due to the highly motivating nature of such an atmosphere and the negative effects extrinsic rewards can have on participant's intrinsic motivation. The results also lend support to Harter's (1974; 1978b) findings that when children are "optimally challenged", they are more likely to experience enjoyment.

Rogers and Ponish (1987) examined the impact that giving children choice had on the level of task difficulty chosen. A total of 80 Grade 1 students took part in a game whereby they had to throw a beanbag into a target set at various distances. One group of children was told where to stand while the other group were allowed to stand anywhere they wanted. The findings demonstrated that children in the no-choice group persisted longer than the choice group at the closer distances because they experienced more success. Those in the choice group persisted longer than those who had no choice and were assigned to the hard condition. In fact, children were more likely to tolerate some level of self-imposed failure in order to attempt a higher level of challenge. However, in order to keep persisting, the children had to start experiencing some level of success, or else they would just quit and/or move on to a new distance.

The theoretical framework from which optimal challenge has been framed suggests that both the perceived challenge of the activity and the perceived ability level of the individual must be taken into consideration to create activities in which individuals feel optimally challenged (Dobos, 1996). Because of the close theoretical link the operational definition of flow has to optimal challenge (i.e., balance between skill and challenge), several studies have used the definition of flow as a measure of optimal challenge. According to the model, the most enjoyable subjective experiences occur when individuals report high levels of perceived challenge and skill. These experiences have been operationalised as being in a flow state. When perceived challenge and skill are below average, participants are said to experience apathy. However, when the challenge of the activity exceeds the skill or ability level of the participant, they are said to be in a state of anxiety. The opposite relationship (i.e., high skill, low challenge) is characteristic of a boredom state. This method of operationalising participants' flow states or optimal experiences has been used in a number of studies examining the psychological outcomes of each type of state. As has been suggested previously, the operational definition of a flow state can be used as a starting point to operationalise optimal challenge due to the similarities that this measure has to the way optimal challenge is defined theoretically.

To date, there has been limited research conducted with children in physical activity environments that uses this ratio of perceived challenge and skill as a measure of optimal challenge. Mandigo and Couture (1996) asked children ages nine to fourteen to rate their perceived skill level, challenge of the activity, and fun immediately following their participation in six activities in their physical education class. After coding participants according to the ratio of their perceived skill and challenge values, the children were more likely to indicate high levels of fun when they reported above average levels of skill and challenge. However, children still rated their level of fun high when coded in the boredom quadrant (i.e., high skill, and low challenge). Mandigo, Thompson, and Couture (1998) reported data from two studies that examined the quality of children's physical activity experiences during two different summer physical activity camps. In both studies, they found the flow state quadrant to be characterised by high quality of experience levels (i.e., high levels of fun, intrinsic motivation, positive affect, perceived success and low levels of state anxiety). However, the boredom quadrant was again characterised by high levels of subjective experience. This raises some concerns in regards to how the balance between skill and challenge have been measured.

Turner. Parkes, Cox, and Meyer (1995) found similar results with a group of eight children in a Grade 5 literacy class. Immediately following five pre-selected classes, students would complete a short questionnaire measuring their perceived skill and challenge. affect. activation. cognitive efficiency, degree of engagement, and intrinsic motivation. Results showed that when perceived challenge and skill were equal and above average, students reported higher levels of affect, activation, cognitive efficiency, degree of engagement, and intrinsic motivation. In follow-up interviews, they indicated that when their skill level did not match the challenge of the activity, they would either make the task easier or more difficult to enhance enjoyment and engagement. In examining the characteristics of tasks that facilitated flow experiences, activities that provided students with a chance to modify and have some control were more likely to produce flow-like experiences.

Rowley (1996) took a qualitative approach to examine children's physical activity experiences. Based on interviews, focus groups, field note observations, and journal entries by the children, one of the key findings she reported was that optimally challenging activities (i.e.. appropriate challenges) were viewed by all of the children as being the most fun. Alternatively. inappropriate challenges were viewed quite negatively. Most participants indicated that they did not like activities that were not challenging or when they were too challenging. As a result. many of the children in her study indicated that they actively sought out activities that challenged them in an appropriate manner.

Sanders and Graham (1995) also took a qualitative approach to their study by interviewing and observing two lesser skilled and two higher skilled kindergarten students with respect to the quality of their experiences during physical education class. The conclusion reached by these authors was that children often resorted to "play" to get into situations where they could match their perceived abilities with the perceived challenge of the activities because they were the most enjoyable. This either involved participating in activities that were more challenging than what the teacher was presenting or participating in activities that were less challenging. The consequence of this was that in most of the teacher designed tasks, the children were not in a state of flow because their perceived abilities were unbalanced with the perceived challenge of the activity. Therefore, the authors stressed the importance of teachers being aware of the developmental characteristics of children in order to structure optimally challenging activities for all the children in their classrooms and to involve the children in decision making processes during class time.

Chalip, Csikszentmihalyi, Kleiber, and Larson (1984) published one of the first papers examining adolescents' flow experiences across different physical activity environments. To collect participants' responses, they used the Experience Sampling Method (ESM) which measured the subjective experiences of 75 high-school students over the course of one week. Their results suggested that during informal sports where adult control was non-existent, participants were more likely to report a balance of perceived challenge and skill than in environments that were predominately adult controlled (e.g., organised sports, physical education class). Other studies have also demonstrated perceived control to be an important variable in creating flow experiences (Csikszentmihalyi & Larson, 1984; Kleiber, Larson, & Csikszentmihalyi, 1986).

Consistency across the life-span with respect to the psychological and behavioural outcomes of being optimally challenged has been reported. When using flow state experiences as a theoretical construct to measure optimal challenge, children, adolescents, adults, and seniors have all reported increased adherence to activities, high levels of perceived success, self-esteem, positive affect, concentration, enjoyment, interest, happiness, creativity, motivation, satisfaction, clear feedback, perceived control, relaxation, and altered sense of time when skills and challenge are balanced (e.g., Csikszentmihalyi, Rathunde, & Whalen, 1993; Haworth and Evans, 1995; Jackson, 1992; 1996; Lefevre, 1988; Mannell, Zuzanek, & Larson, 1988; Stein et al., 1995).

Along with examining the impact of optimal challenge on psychological outcomes and situational variables on creating optimally challenging environments, some researchers have turned their attention to investigating the impact of person variables on creating optimally challenging conditions. For example, Meyer, Turner, and Spencer (1997) examined the differences in acceptance of failure and adaptive learning between challenge seekers and challenge avoiders in Grades 5 and 6 math. Although the sample size was rather small ($n = 14$) for powerful statistical analyses, the patterns suggested that the challenge seekers reported a higher preference for task difficulty and taking action in their math studies, had a higher level of positive affect after failure, higher levels of math self-efficacy, were more mastery focused and used deeper strategies than students classified as challenge avoiders.

Children's goal orientation and perceived ability have also been suggested to have an impact on their preference for challenging activities. Sarrazin and Famose (1999) reported that participants' choices of a climbing wall was influenced by their goal orientations and perceived ability. Participants who were ego-oriented but had low levels of perceived climbing ability, tended to choose easy courses. About 25% of these participants, however, tended to choose the very difficult course so that they would fail on purpose and hence, attribute their failure to the difficulty of the wall rather than their skill. For those who were ego-oriented and had high levels of perceived climbing ability, they were more likely to choose the moderate over the difficult course. Participants who were high on task-orientation and low on perceived ability tended to choose the moderate course while those high on both task-orientation and perceived climbing ability exclusively chose the difficult courses. None of the task-oriented participants chose an easy course with a high probability of success.

Boggiano, Main, and Katz (1988) investigated the impact that various dispositional and situational variables had on children's preference for challenging tasks. In their first study, Boggiano et al. (1988) found that children's self-reported perceptions of academic competence and personal control were related positively to intrinsic interest in schoolwork and preference for challenging school activities. In the second study, the authors reported that children with high levels of academic competence and personal control were more likely to report higher preferences for challenging activities when placed in evaluative, and controlling conditions than those with low levels of academic competence and personal control. However, when the controlling condition did not exist, there were no differences between those with high and low levels of academic competence and personal control.

Despite these findings, the high challenge and high skill ratio or "flow" quadrant has not always been reported as the most preferred state. Haworth (1993) reported that in a sample of British college students, the control channel (i.e., moderate challenge, and high skill) was the preferred channel as it was associated with high levels of enjoyment, interest, relaxation, and happiness. Mandigo et al. (1998) and Turner et al. (1995) have also found the boredom quadrant (i.e., high skill, and low challenge) to be associated with positive subjective states. Although in all of these studies the subjective state in other channels or quadrants was not significantly higher than the flow quadrant, it does demonstrate that the high challenge/ high skill relationship does not always produce the highest subjective state. The results from these studies demonstrate the importance of examining optimal challenge from an individual level and to determine the impact of dispositional and situational characteristics on creating optimally challenging environments.

Limitations of Current Optimal Challenge Research

To date, no one theory adequately addresses the construct of optimal challenge. Many different theories have pointed to its importance (e.g., Csikszentmihalyi, 1975; Deci & Ryan, 1985; Harter, 1978a), but they have not fully explored the conditions that create the psychological and behavioural outcomes of optimal challenge. Quite often, optimal challenge has been identified as an antecedent to other constructs such as intrinsic motivation (Ryan & Deci, 2000) and perceived competence (Harter, 1992), but the theories do not fully explain how to facilitate optimal challenge experiences.

One of the potential reasons for this paucity of empirical evidence is the lack of instruments with good psychometric properties to tap into children's optimal challenge experiences during physical activities. In a review of instruments used within the sport psychology field, Vallerand and Fortier (1998) reviewed eight commonly used instruments to measure intrinsic motivation. Of these questionnaires, only the Motivational Orientations in Sports Scale (Weiss et al., 1985) had a challenge subscale and this was intended to measure a trait. The Experience Sampling Form (ESF) commonly used to measure the flow states and optimal challenge experiences of individuals soon after participating in an activity (Csikszentmihalyi & Csikszentmihalyi, 1988) was not included in their review. The ESF provides a basic and unidimensional measure of optimal challenge by asking participants to report their perceived skill level and challenge of the activity soon after they had just participated. However, some have questioned the validity of interpretations made from this instrument. For example, concerns over using only two items to measure optimal challenge (Ellis & Voelkl, 1998), the common procedure of standardising individual challenge and skill responses (Ellis, Voelkl, & Morris, 1994), the one dimensional perspective of skill and challenge (Voelkl & Ellis, 1998), and the validity of this measure with children (Mandigo & Thompson, 1998) have come into question.

Continuing the Research on Optimal Challenge as a Construct

The previous review of literature demonstrates the importance of optimal challenge as a psychological construct housed within humanistic theories of intrinsic motivation. Despite its importance on facilitating positive affective outcomes (e.g., enjoyment, competence), it has not been adequately researched in terms of its impact on children's motivation to be active. This is partly due to the absence of psychometrically sound instruments that are capable of tapping into this construct. Therefore, if further research into optimal challenge is to take place, ways to examine this construct are needed. However, before pursuing down this road, it is important to identify one's own philosophical assumptions when it comes to developing and answering research questions.

Identifying Assumptions of an Organismic Approach

Slife (1998) suggests that all researchers should identify their philosophical assumptions towards science and to provide a rationale for choosing a particular method. This is especially important in the study of human motivation that has been examined from many different philosophical perspectives. Originally, the study of human motivation was influenced by the natural sciences that took a mechanistic perspective. This approach to human motivation was eventually challenged with the introduction of organismic ways of thinking (e.g., White, 1959). According to organismic theories of motivation, humans are active participants in their environments and are influenced by their experiences with the environment (Deci, 1975). Out of organismic approaches to motivation emerged social-cognitive and humanistic theories of motivation. What both of these approaches recognised is that humans are influenced by their cognition and that each person's experiences within the environment has a different impact on his/her cognition. The optimal challenge construct presented in this dissertation adopts an organismic approach pertaining to assumptions taken from humanistic and social-cognitive theories.

At the person level, the construct of optimal challenge takes on a humanistic perspective to motivation in the sense that it recognises that each individual has a desire to better themselves and is based on an individual's direct experience (Deci, 1995). Humanism has been defined as: "A theory ... based on the autonomy of humans but allowing for the constraints of biology and social structures in which humans find themselves" (Slife & Williams, 1995, p. 231). Humanistic theories propose that individuals have a desire to achieve self-actualisation and are self-determined in their quest to reach this state. Such approaches defy the behaviourist and psychoanalytic approaches towards motivation because they propose that humans have the freedom to choose their course of action and are not solely constrained by biological forces (Slife & Williams, 1995). One of the assumptions embedded within the three intrinsic motivation theories presented in this dissertation is that individuals have a desire to be optimally challenged. This is due to their internal motivation to enhance their competence and to experience a positive subjective state. Hence, it is this desire to strive towards self-actualisation that is embedded within the concept of optimal challenge. The humanistic approach fits nicely with the concept of intrinsic motivation in that the behaviours that individuals engage in are done so to fulfil their own personal needs as opposed to the needs or constraints of others. As outlined by Deci and Ryan (1994), such motives are closely related to intrinsic motivation. Hence, from a humanistic perspective, it is the perception of one's abilities and the subjective experiences that contribute to one's intrinsic motivation to be optimally challenged (Dobos, 1996).

Based on the theoretical frameworks outlined to date and the empirical examples related to children's desire to challenge themselves, the assumption adopted in this dissertation is that children do experience optimal challenge during their physical activities. This assumption is embedded within the three theories of motivation detailed previously, limited empirical evidence, and on personal observations when working with children. In combination, these existing pieces of evidence suggest that the construct does exist. However, such assumptions need to be tested. These assumptions need to be put to the test through methodologies that attempt to control for researcher bias and which use developmentally appropriate instruments that are able to capture the existence of such constructs. Given that the construct of optimal challenge has been relatively understudied and there does not exist a psychometrically sound instrument that is able to tap into children's optimal challenge experiences during physical activity, there is a need to provide various sources of construct validation evidence in order to create a developmentally appropriate instrument that will permit valid interpretations of children's optimal challenge experiences. Therefore, the importance of validity is central in creating a developmentally appropriate measure of optimal challenge experiences.

Validity From a Unitary Perspective

The issue of validity is embedded within any study that attempts to measure a construct established inside a theoretical framework. Although many different definitions of validity have been put forth, one of the most commonly referenced definitions is: " ... an integrated evaluative judgement of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores and other models of assessment" (Messick, 1989, p. 13). Evidence of validity is important in order to gain confidence that the construct identified within a theoretical framework is truly being measured by the instrument(s) being used. Since a construct itself is not directly observable, it is best to obtain various sources of evidence to support its existence. In such instances, the construct must be operationalised, and related to measures of other constructs embedded within the theoretical framework supported with measures of specific real-world criteria (Crocker & Algina, 1986). Given that validity is an ongoing process, there are several possible methods to provide validity evidence. The traditional approaches include content validity, predictive validity, concurrent validity, criterion-related validity, and construct validity. Recent thinking pertaining to sources of validity evidence within the behavioural sciences has suggested a unitary perspective to validity and that all sources of validity are subsumed by construct validity (e.g., Fox, 1998; Messick, 1989). The studies outlined in this dissertation adopt this perspective and put forth various pieces of evidence that help to establish the construct validity of a self-report measure of children's optimal challenge experiences.

Construct validity also has been defined as:

> "... any evidence that bears on the interpretation or meaning of the test scores ... Almost any kind of information about a test can contribute to an understanding of its construct validity, but the contribution becomes stronger if the degree of fit of the information with the theoretical rationale underlying score interpretation is explicitly evaluated." (Messick, 1989, p. 17).

Hence, providing evidence of construct validity is not *only* an issue of putting forth evidence of the psychometric properties of an instrument, but linking inferences back to a theoretical framework and real life experiences from which inferences about the construct are drawn (Flattery, 1990).

Construct validation evidence can be provided in a number of ways. Through judgmental and logical analysis, the content-relevance and content-representativeness of the items developed to measure the construct can be determined. Correlational studies can be used to examine both the internal (e.g., factor structure) and external (e.g., relationship to other constructs) structures of the new instrument. As well, experimental studies that manipulate independent variables can be used to see if the instrument produces results that are consistent (i.e., either convergent or divergent) with theoretical predictions related to the construct in question. These methodologies (i.e., logical, correlational, and experimental) will be used in the studies outlined in this dissertation to provide construct validation evidence pertaining to a measure of children's optimal challenge experiences in physical activity settings.

Developmental Considerations for Constructing Self-Reports for Children

Based on current knowledge of child development, research should consider the ability of children to evaluate their subjective experiences pertaining to an abstract concept (Brustad, 1998). Creating a developmentally appropriate instrument that captures children's psychological experiences is not an easy task. Simply changing the wording of questionnaires intended for adults is not good enough (Scott, 2000). Researchers must not only ensure that the wording is appropriate for the children, but that the construct being measured is actually experienced by the participants and that it can be operationalised and measured by an appropriate instrument. Brustad (1998) and others have identified several guidelines and considerations that researchers should follow when designing and/or choosing developmentally appropriate instruments for children.

First, the ability for children to use abstract reasoning must be taken into consideration. Based on Piaget's theory of cognitive development, children's ability to think in abstract terms does not develop until they reach the concrete operations stage (Crain, 1992). The onset of this ability usually occurs at around age seven when the child achieves conservation. Conservation is characterised by children's ability to focus on more than one perceptual dimension in which they use logical operations (Piaget, 1970). Children in the pre-operational stage of cognitive development often have difficulty viewing the world from another person's perspective apart from their own. That is to say, they are egocentric (Shaffer, 1989). Once children have achieved conservation, they become less egocentric and are able to view the world from different perspectives. However, at the beginning of the concrete operational stage, children need concrete examples to effectively solve problems. Although they have the ability to reason in abstract terms, this ability is still developing. Their ability to reason, however, is enhanced when children take part in ecologically valid tasks (Woodhead & Faulkner, 2000). Hence, measures of children's psychological experiences need to be framed in such a way that they represent a concrete and realistic example as opposed to abstract concepts (Brustad, 1998).

Secondly, children's ability to process information accurately must be considered. Simply stated, an individual receives input from the environment, processes the information using various cognitive functions, and selects an appropriate response to the stimulus. Through normal development, children acquire the ability to process information using more complex cognitive skills. Their ability to problem solve, focus attention on a problem, and retain information are all enhanced with cognitive development (Miller, 1993). However, because children's ability to process information accurately and efficiently is still developing, measures that require a high level of information-processing ability should not be used until later childhood.

Third, the structure of a participant's self-system must be considered. By around seven or eight years of age, children are able to judge their competencies within different domains such as academic competence, athletic competence, peer acceptance, behavioural conduct, and physical appearance (Harter, 1990). Children's perceptions of their abilities also change over time. Perceptions of ability for children between five and nine years of age focus on effort expended, personal mastery, and social reinforcement to form impressions of their abilities. With age, children are better able to evaluate effort, mastery and social reinforcement and begin to differentiate between performance, effort, and ability (Stipek, 1998). Because of these concerns, measures must reflect the developmental level of the individual with respect to their self-system.

Finally, the use of children's language is vital in ensuring that the items accurately reflect a relevant construct. If the language that is used is too difficult for children to understand or if it is not reflective of how children would describe the construct, then the instrument may contain a source of construct-irrelevant difficulty (Messick, 1995). For children. the basic dimensions of language development are established by around five years of age (Stone & Lemanek, 1990). At around the same time, children are able to provide a verbal description of tasks (Singer & Revenson, 1996). Therefore, in order to capitalise on the development of language skills and the ability to verbalise tasks, researchers using a self-report instrument must ensure it is at the appropriate reading level yet still reflective of how children would describe a particular construct.

Outline of Studies to Examine Optimal Challenge with Children.

As suggested by current research and theory, the construct of optimal challenge plays a major role in motivating children during physical activities. Despite this importance, there currently does not exist a measurement instrument with sufficient validity evidence that can provide information on children's optimal challenge during physical activities. To date, the best measure of optimal challenge according to Deci and Ryan (1985) is Csikszentmihalyi's (1975) operationalisation of flow. However, as indicated previously, this measure has yet to be validated with children in a physical activity environment and has been the subject of criticism by many regarding the interpretations that result from this instrument (Ellis et al., 1994; Voelkl & Ellis, 1998). The purpose of the series of studies outlined in this dissertation is to create developmentally appropriate items for a self-report instrument of optimal challenge for use with children and to present various sources of validity evidence that will increase the likelihood that the inferences made from such an instrument reflect children's perceptions of being optimally challenged.

Through a series of studies, a number of sources of construct validation evidence are examined with the intention of documenting whether optimal challenge is a part of children's real world experiences and to provide evidence that a self-report instrument is able to tap into these experiences in physical activity settings. Each study is progressive in nature. In order to determine what children's perceptions of optimal challenge were and to use their own words to guide the development of appropriate items, 33 children were interviewed in study one about their optimal challenge and non-optimal challenge experiences during physical activity. Based on these interviews and an extensive literature review, items for a self-report instrument were generated. In the second study, expert judges rated the relevance of these items as they pertained to their match of children's perceptions of physical activities where their skills were higher than the challenge, the challenge was higher than their skills or where they perceived the challenge to be equal to their skill level. Based on the judges' ratings and comments, the third study asked children how similar or dissimilar they felt the items were to one another. Participants were also asked to group similar items together and to name the groupings of items they had generated. The final study in this dissertation took the remaining items and placed them into a self-report format where children rated the degree to which they felt their skills and challenge were higher than each other or equal to one another. The instrument was then field-tested and various sources of validity (e.g., predictive, concurrent, inter-correlations) were used to provide construct validity evidence. These studies intend to serve as the first steps of a research program that will eventually lead to a better understanding of how to create optimally challenging experiences for children and what the outcomes of such experiences may be. However, before this can take place, a way of measuring optimal challenge experiences is required.

References

Anshel, M. H., Weinberg, R., & Jackson, A. (1992). The effect of goal difficulty and task complexity on intrinsic motivation and motor performance. Journal of Sport Behaviour. 15, 159 - 176.

Biddle, S. J. H. (1997). Cognitive theories of motivation and the physical self. In K. Fox (Ed.), The physical self (pp. 59 - 82). Champaign, IL: Human Kinetics.

Boggiano, A. K., Main, D. S., & Katz, P. A. (1988). Children's preferences for challenge: The role of perceived competence and control. Journal of Personality and Social Psychology 54, 134 - 141.

Brustad, R. J. (1998). Developmental considerations in sport and exercise psychology measurement. In J. Duda (Ed.), Advances in sport and exercise psychology measurement (pp. 461 - 470). Morgantown, WV: Fitness Information Technology.

Buschner, C. (1996). Teaching children movement concepts and skills. Champaign, IL: Human Kinetics.

Chalip, L., Csikszentmihalyi, M., Kleiber, D., & Larson, R. (1984). Variations of experience in formal and informal sport. Research Quarterly, 55, 109 - 116.

Crain, W. (1992). Theories of development (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Philadelphia, PA: Harcourt Brace Jovanovich.

Csikszentmihalyi, M. (1975). Beyond boredom and anxiety. San Francisco, CA: Jossey-Bass.

Csikszentmihalyi, M. (1988a). The flow experience and its significance for human psychology. In M. Csikszentmihalyi & I. Csikszentmihalyi (Eds.), Optimal experience (pp. 15 - 35). Cambridge, UK: Cambridge University Press.

Csikszentmihalyi, M. (1988b). The future of flow. In M. Csikszentmihalyi & I. Csikszentmihalyi (Eds.), Optimal experience (pp. 364 - 383). Cambridge, UK: Cambridge University Press.

Csikszentmihalyi, M. (1990a). Flow. New York: Harper & Row.

Csikszentmihalyi, M. (1990b). What good are sports? Reflections on the psychological outcomes of physical performance. New Zealand Journal of Health, Physical Education, and Recreation, 23(2), 2 - 7; 10 - 11.

Csikszentmihalyi, M. (1992). A response to the Kimiecik & Stein and Jackson papers. Journal of Applied Sport Psychology, 4, 181 - 183.

Csikszentmihalyi, M. & Csikszentmihalyi, I. (1988). Optimal experience. Cambridge, UK: Cambridge University Press.

Csikszentmihalyi, M., & Larson, R. (1984). Being adolescent. New York: Basic Books.

Csikszentmihalyi, M. & Larson, R. (1987). Validity and reliability of the experience-sampling method. The Journal of Nervous and Mental Disease, 175, 526 - 536.

Csikszentmihalyi, M., Rathunde, K. & Whalen, S. (1993). Talented teenagers. The roots of success and failure. Cambridge, UK: Cambridge University Press.

Danner, F. W. & Lonky, E. (1981). A cognitive-developmental approach to the effects of rewards on intrinsic motivation. Child Development, 52, 1043 - 1052.

Deci, E. L. (1975). Intrinsic motivation. New York: Plenum.

Deci, E. L., & Flaste, R. (1995). Why we do what we do. New York: Penguin.

Deci, E. L. & Olson, B.C. (1989). Motivation and competition: Their role in sports. In J. H. Goldstein (Ed.), Sports, games and play: Social and psychological viewpoints (pp. 83 - 110). Hillsdale, NJ: Erlbaum.

Deci, E. L. & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behaviour. New York: Plenum.

Deci, E. L. & Ryan, R. M. (1994). Promoting self-determined education. Scandinavian Journal of Educational Research, 38, 3-14.

Dobos, J. A. (1996). Collaborative learning: Effects of student expectations and communication apprehension on student motivation. Communication Education, 45, 118 - 134.

Duda, J. L., Chi, L., Newton, M. L., Walling, M. D., & Catley, D. (1995). Task and ego-orientation and intrinsic motivation in sport. International Journal of Sport Psychology, 26, 40 - 63.

Ellis, G., Voelkl, J. E., & Morris, C. (1994). Measurement and analysis issues with explanations of variance in daily experiences using the flow model. Journal of Leisure Research, 26, 337 - 356.

Flattery, R. C. (1990). Methodological and psychometric considerations in child reports. In A. M. La Greca (Ed.), Through the eyes of the child (pp. 57 - 108). Needham Heights, CA: Allyn & Bacon.

Fox, K. R. (1998). Advances in the measurement of the physical self. In J. Duda (Ed.), Advances in sport and exercise psychology measurement (pp. 295 - 310). Morgantown, WV: Fitness Information Technology.

Frederick, C. M. & Ryan, R. M. (1995). Self-determination in sport: A review using cognitive evaluation theory. International Journal of Sport Psychology, 26, 5 - 23.

Goudas, M., Biddle, S., Fox, K., & Underwood, M. (1995). It ain't what you do, it's the way that you do it! Teaching style affects children's motivation in track and field lessons. The Sport Psychologist, 9, 254 - 264.

Graham, G. (1992). Teaching children physical education. Champaign, IL: Human Kinetics.

Harter, S. (1974). Pleasure derived by children from cognitive challenge and mastery. Child Development, 45, 661 - 669

Harter, S. (1978a). Effectance motivation reconsidered. Human Development, 21, 34 - 64.

Harter, S. (1978b). Pleasure derived from challenge and the effects of receiving grades on children's difficulty level choices. Child Development, 49, 788 - 799.

Harter, S. (1981a). A model of mastery motivation in children: Individual differences and developmental change. Aspects of the development of competence: The Minnesota symposia on child psychology: Vol. 14, (pp. 215 - 255). Hillsdale, NJ: Lawrence Erlbaum Associates.

Harter, S. (1981b). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. Developmental Psychology, 17, 300 - 312.

Harter, S. (1985). Manual for the self-perception profile for children. Denver, CO: University of Denver Press.

Harter, S. (1990). Causes, correlations and the functional role of global self-worth: A life-span perspective. In R. J. Sternberg, & J. Kalligan (Eds.), Competence considered (pp. 67 - 97). London, UK: University Press.

Harter, S. (1992). The relationship between perceived competence, affect, and motivational orientation within the classroom: Processes and patterns of change. In A. K. Boggiano & T. S. Pittman (Eds.), Achievement and motivation: A social-developmental perspective (pp. 77 - 114). Cambridge, UK: Cambridge University Press.

Harter, S., & Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for young children. Child Development, 55, 1969 - 1982.

Haworth, J. T. (1993). Skill-challenge relationships and psychological well-being in everyday life. Society and Leisure, 16, 115 - 128.

Haworth, J. & Evans, S. (1995). Challenge, skill and positive subjective states in the daily life of a sample of YTS students. Journal of Occupational and Organizational Psychology, 68, 109 - 121.

Jackson, S. (1992). Athletes in flow: A qualitative investigation of flow states in elite figure skaters. Journal of Applied Sport Psychology, 4, 161 - 180.

Jackson, S. A. (1995). The growth of qualitative research in sport psychology. In T. Morris & J. Summers (Eds.), Sport psychology. Theory, applications and issues (pp. 575 - 591). New York: John Wiley & Sons.

Jackson, S. A. (1996). Toward a conceptual understanding of flow experience in elite athletes. Research Quarterly for Exercise and Sport, 67, 76 - 90.

Jackson, S. A., & Roberts, G. C. (1992). Positive performance states of athletes: Toward a conceptual understanding of peak performance. The Sport Psychologist, 6, 156 - 171.

Kimiecik, J. C., & Stein, G. L. (1992). Examining flow experiences in sport contexts: Conceptual issues and methodological concerns. Journal of Applied Sport Psychology, 4, 144 - 160.

Kleiber, D. A. (1981). Searching for enjoyment in children's sports. The Physical Educator, 38, 77 - 84.

Kleiber, D., Larson, R., & Csikszentmihalyi, M. (1986). The experience of leisure adolescence. Journal of Leisure Research, 18, 169 - 176.

Klint, K.A., & Weiss, M. R. (1987). Perceived competence and motives for participating in youth sports: A test of Harter's competence motivation theory. Journal of Sport Psychology, 9, 55 - 65.

Kowal, J., & Fortier, M. S. (1999). Motivational determinants of flow: Contributions from self-determination theory. The Journal of Social Psychology, 139, 355 - 368.

Lefevre, J. (1988). Flow and the quality of experiences during work and leisure. In M. Csikszentmihalyi & I. Csikszentmihalyi (Eds.), Optimal experience (pp. 307 - 318). Cambridge, UK: Cambridge University Press.

Mandigo, J. L., & Couture, R. T. (1996). An overview of the components of fun in physical education, organised sports and physical activity programs. AVANTE, 2, 56 - 72.

Mandigo, J. L., & Thompson, L. P. (1998). Go with their flow: How flow theory can help practitioners to intrinsically motivate children to be physically active. The Physical Educator, 55, 145 - 159.

Mandigo, J. L., Thompson, L. P., & Couture, R. T. (1998). Equating flow theory with the quality of children's physical activity experiences. Journal of Sport and Exercise Psychology, 20(Suppl.), S94.

Mannell, R., Zuzanek, J., & Larson, R. (1988). Leisure states and "flow" experiences: Testing perceived freedom and intrinsic motivation. Journal of Leisure Research, 20, 289 - 304.

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. Research Quarterly for Exercise and Sport, 60, 48 - 58.

McAuley, E., Wraith, S., & Duncan, T. E. (1991). Self-efficacy, perceptions of success, and intrinsic motivation for exercise. Journal of Applied Social Psychology, 21, 139 - 155.

McInnman, A. D., & Grove, J. R. (1991). Peak moments in sport: A literature review. Quest, 43, 333 - 351.

McKiddie, B., & Maynard, I.. W. (1997). Perceived competence of schoolchildren in physical education. Journal of Teaching in Physical Education, 16, 324 - 339.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed.) (pp. 13 - 103). New York: American Council on Education.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist, 50, 741 - 749

Meyer, D. K., Turner, J. C., & Spencer, C. A. (1997). Challenge in a mathematics classroom: Students' motivation and strategies in project-based learning. The Elementary School Journal, 97, 501 - 521.

Miller, P. H. (1993). Theories of developmental psychology (3rd ed.). New York: W.H. Freeman and Company.

Morris, G. S. D., & Stiehl, J. (1999). Changing kids' games (2nd ed.). Champaign, IL: Human Kinetics.

National Association for Sport and Physical Education (2000). Appropriate practices for elementary school physical education. Reston, VA: NASPSE Publications.

Ommundsen, Y., & Vaglum, P. (1991). Soccer competition anxiety and enjoyment in young boy players. The influence of perceived competence and significant others' emotional involvement. International Journal of Sport Psychology, 22, 35 - 49.

Orlick, T. D., & Mosher, R. (1978). Extrinsic awards and participant motivation in a sport related task. International Journal of Sport Psychology, 9, 27 - 39.

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), Carmichael's manual of child psychology (pp. 703 - 732). New York: Wiley.

Privette, G. (1983). Peak experiences, peak performance, and flow: A comparative analysis of positive human experiences. Journal of Personality and Social Psychology, 45, 1361 - 1368.

Rea, D.W. (2000). Optimal motivation for talent development. Journal for the Education of the Gifted, 23, 187 - 216.

Reeve, J. (1996). Motivating others. Boston, MA: Allyn and Bacon.

Rogers, C. S., & Ponish, K. K. (1987, April). Control of level of challenge: Effects on intrinsic motivation to play. Paper presented at the Society for Research in Child Development, Baltimore, MA.

Rowley, D. A. (1996). Physical activity through the eyes of six ten-year-old children: A case study. Unpublished doctoral dissertation. University of Alberta, Edmonton, Alberta, Canada.

Ryan, R. M., Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemporary Educational Psychology, 25, 54 - 67.

Ryan, R. M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. Journal of Personality and Social Psychology, 45, 736 - 750.

Sanders, S., & Graham, G. (1995). Kindergarten children's initial experiences in physical education: The relentless persistence for play classes with the zone of acceptable responses. Journal of Teaching in Physical Education, 14, 372 - 383.

Sarrazin, P., & Famose, J. (1999). Children's goals and motivation in physical education. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand, & R. Seiler (Eds.), Psychology for physical educators (pp. 27 - 50). Champaign, IL: Human Kinetics.

Scanlan, T. K., & Simons, J. P. (1992). The construct of sport enjoyment. In G. C. Roberts (Ed.), Motivation in sport and exercise (pp. 199 - 215). Champaign, IL: Human Kinetics.

Scott, J. (2000). Children as respondents. The challenge for quantitative methods. In P. Christensen, & A. James (Eds.), Research with children: Perspectives and practices (pp. 98 - 119). London: Falmer.

Shaffer, P. S. (1989). Developmental psychology. Pacific Grove, CA: Brooks/Cole.

Singer, R. N. (1996). Moving toward the quality of life. Quest, 48, 246 - 252.

Singer, D. G., & Revenson, T. A. (1996). A Piaget primer: How a child thinks. New York: Plume.

Slife, B. D. (1998). Raising the consciousness of researchers: Hidden assumptions in the behavioural sciences. Adapted Physical Activity Quarterly, 15, 208 - 221.

Slife, B. D., & Williams, R. N. (1995). What's behind the research? Discovering hidden assumptions in the behavioural sciences. Thousand Oaks, CA: Sage.

Stein, G. L., Kimiecik, J. C., Daniels, J., & Jackson, S. A. (1995). Psychological antecedents of flow in recreational sport. Personality and Social Psychology Bulletin, 21, 125 - 135.

Stipek, D. (1998). Motivation to learn. From theory to practice. Boston, MA: Allyn & Bacon.

Turner, J. C., Parkes, J., Cox, K. E., & Meyer, D. K. (1995). The role of optimal challenge in students literacy engagement. In K. A. Hinchman, D. J. Leu, & C. K. Kinzer (Eds.), Perspectives on literacy research and practice. Forty-fourth yearbook of the National Reading Conference (pp. 126 - 136).

Vallerand, R. J. (1983). The effect of differential amounts of positive verbal feedback on the intrinsic motivation of male hockey players. Journal of Sport Psychology, 5, 100 - 107.

Vallerand, R. J., & Fortier, M. S. (1998). Measures of intrinsic and extrinsic motivation in sport and physical activity: A review and critique. In J. Duda (Ed.), Advances in sport and exercise psychology measurement (pp. 81 - 101). Morgantown, WV: Fitness Information Technology.

Vallerand, R. J., Gauvin, L. I., & Halliwell, W. R. (1986). Effects of zero - sum competition on children's intrinsic motivation and perceived competence. The Journal of Social Psychology, 126, 465 - 472.

Vallerand, R. J., & Reid, G. (1984). On the causal effects of perceived competence on intrinsic motivation: A test of cognitive evaluation theory. Journal of Sport Psychology, 6, 94 - 102.

Voelkl, J. E., & Ellis, G. D. (1998). Measuring flow experiences in daily life: An examination of the items used to measure challenge and skill. Journal of Leisure Research, 30, 380 - 389.

Wankel, L. M. (1993). The importance of enjoyment to exercise adherence and psychological benefits from physical activity. International Journal of Sport Psychology, 24, 151 - 169.

Weiss, M. R. (1986). Self-esteem and achievement in children's sport and physical activity. In D. Gould, & M. Weiss (Eds.), Advances in pediatric sport sciences:: Vol. 2, (pp. 87 - 119). Champaign, IL: Human Kinetics.

Weiss, M. R., Bredemeier, B. J., & Shewchuk, R. M. (1985). An intrinsic/extrinsic motivation scale for the youth sport setting: A confirmation factor analysis. Journal of Sport Psychology, 7, 75 - 91.

Weiss, M. R., & Bressan, E. S. (1985). Relating instructional theory to children's psychosocial development. Journal of Physical Education, Recreation and Dance, November/December, 34 - 36.

White, R. W. (1959). Motivation reconsidered: The concept of competence. Psychological Review, 66, 297 - 333.

Whitehead, J. R., & Corbin, C. B. (1991). Youth fitness testing: The effect of percentile-based evaluative feedback on intrinsic motivation. Research Quarterly for Exercise and Sport, 62, 225 - 231.

Woodhead, M., & Faulkner, D. (2000). Subjects, objects or participants? Dilemmas of psychological research with children. In P. Christensen, & A. James (Eds.), Research with children: Perspectives and practices (pp. 9 - 35). London: Falmer.

CHAPTER 2

Study 1: Listening to the Voices of Children: Taking the First Steps in the Development of a Self-Report Instrument of Children's Perceptions of Optimal Challenge

In an era where concerns regarding the physical activity levels of children have been raised (e.g., Andersen, 2000; Craig, Russell, Cameron, & Beaulieu, 1999; Secretary of Health and Human Services and the Secretary of Education, 2000; Tremblay & Wilms, 2000), an increased emphasis has been placed on understanding how to motivate children to be physically active. To date, research has provided us with a fairly good understanding of what motivates children (e.g., Vanden Auweele, Baker, Biddle, Duard, & Seiler, 1999; Roberts, 1992; Smoll, Magill, & Ash, 1988), but there is still a lack of understanding regarding how to intrinsically motivate children to be physically active. One of the constructs that may shed some light on how to foster intrinsic motivation is optimal challenge. Individuals are optimally challenged when they perceive the challenge(s) of the activity to be balanced with their abilities to do the task(s) (Csikszentmihalyi, 1990; Reeve, 1996; Weiss, 1986). When children perceive an activity to be too easy or too hard, it may lead to boredom or frustration respectively. However, when children perceive their skills to be equal to the challenge of the activity, then they are likely to be optimally challenged. Many researchers (e.g., Biddle & Chatzisarantis, 1999; Csikszentmihalyi, 1990; Mandigo & Thompson, 1998; Reeve, 1996; Weiss & Bressan, 1985) have hypothesised that when children are optimally challenged, they are intrinsically motivated to do the activity.

Three humanistic theories of intrinsic motivation have integrated the importance of optimal challenge. According to Optimal Experience Theory (Csikszentmihalyi, 1990), Cognitive Evaluation Theory (Deci & Ryan, 1985), and Competence Motivation Theory (Harter, 1978a), optimal challenge is imperative to facilitating intrinsic motivation. When individuals perceive their skills to be equal to the challenge of an activity, they are more likely to have a positive subjective experience (e.g., enjoyment, competence) than if they are unequal. One important feature that must be pointed out is that it is individuals perception of their skills and the perceived degree of challenge that impacts on their subjective experience (Dobos, 1996). An instructor could structure a learning environment which he/she feels is optimal for the learners' abilities, however, the subjective experience individuals have is ultimately influenced by their own perceptions.

Despite the importance that optimal challenge has on children's intrinsic motivation, the only instruments currently available to researchers to study children's optimal challenge are unidimensional (e.g. Harter, 1978a; Delignières, 1999) or intended for adults (e.g., Csikszentmihalyi & Larson, 1987). For example, Harter's (1978b) scale asked participants to rate on a four point Likert scale the perceived difficulty of the task (1 = very easy; 4 = very hard). The scale does not ask participants to indicate whether the activity was optimally challenging for them (e.g., not too easy yet not too hard). Other indicators of optimal challenge have used objective (e.g., task difficulty, distance from target) as opposed to subjective measures (e.g., Danner & Lonky, 1981; Delignières, Famose, Thépaut-Mathieu, & Fleurance, 1993). As indicated previously, not everyone perceives the same level of difficulty the same way and thus, an objective measure may be inadequate due to the subjective nature of the construct. In order to directly assess children's perceptions of optimal challenge as it pertains to the relationship between perceived skill and perceived challenge, a new psychological instrument that is psychometrically sound and sensitive to the subjective nature of optimal challenge for children is needed and warranted.

In order to successfully create developmentally appropriate items that accurately reflect the construct of optimal challenge for children, there is a need to begin at the start of the validation process. Initial steps should focus on content validity so that the subjective nature of optimal challenge as experienced and expressed through the voices of children are the basis for "test construction". By grounding the investigation with children's views, potential problems of invalid interpretations generated from self-report instrument may be controlled for as much as possible.

In essence, validity is: "... the appropriateness, meaningfulness, and usefulness of specific inferences made from test scores" (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1985, p. 9). The importance of content validity is embedded within current thinking of

Messick's (1989) unitary perspective of validity. Messick (1988; 1989) argued that validation of test scores should focus around a unitary perspective that contains empirically grounded interpretations of the construct as sources of validity evidence. This unified representation of validity was aligned under the banner of construct validity which served to bind the validity of the test to the validity of test score interpretations (Messick, 1989). The underlying argument of the unitary perspective is that various sources of evidence (e.g., content validity, criterion-related validity, construct validity) are needed in order to provide support for the validity of an instrument.

In order to demonstrate the trustworthiness of validity interpretations from various sources of evidence, a deliberate and meticulous validation process is needed. Content validation plays an important role in providing evidence of an instrument's construct validity because it is concerned with "... showing how well the content of the test samples the class of situations or subject matter about which conclusions are to be drawn" (Messick, 1989, p. 16). Often, this content is obtained by going to primary sources. These primary sources are those people who experience and can describe the construct in question. Secondary sources of content validity can also be used which may include consulting existing literature on the construct in question in order to gain an understanding of its meaning.

The importance of content validity in providing evidence of construct validity is extremely important when developing self-report instruments. When conducting research with children that uses self-reports to tap into psychological constructs, questions surrounding content validity become extremely important because of the developmental differences that exist between children and adults (Brustad, 1998). Not only must children be capable of understanding the items, but the construct must be defined in a manner that is consistent with the way it is experienced by children. It can not be assumed that just because a construct has been suggested to exist in the "adult world" that the construct is experienced or interpreted in exactly the same way by children. Hence, questions of validity in children's research must begin with an investigation of children's descriptions of how they experience the construct under examination. Once this can be determined, methodologies need to be implemented that will provide various sources of validity evidence pertaining to the inferences drawn from interpretations of scores from self-report instruments. By following this process, it is more likely that the self-report instruments will be developmentally appropriate because they reduce potential sources of error (e.g., task-irrelevance, appropriate language) within the instrument.

It would stand to reason that if one wants to find someone's true score, the amount of measurement error must be reduced as much as possible (Crocker & Angina, 1986). For children, a developmentally appropriate instrument is one that attempts to reduce the amount of measurement error so that inferences made from an individual's observed score are as close as possible to her/his true score for the construct in question. Developmentally appropriate instruments attempt to reduce this error by taking into consideration the developmental characteristics of the child (i.e., their perceptions of the construct). By reducing known sources of error, the observed score is more likely to reflect the true score. For example, construct-irrelevant variance can be attributed to sources of error that is a result of the score being inflated because it was too easy, too difficult, or confounded by other extraneous factors. This is an extremely important issue for those wishing to design and use developmentally appropriate self-report instruments with children. Items need to be designed around the type of language used by participants and current reading level. If the questions are not understood by the participants due to an inability to read and comprehend the item, then scores can not be interpreted with confidence.

In light of the previous arguments, a new self-report instrument to measure optimal challenge during physical activity is needed if it is to be used effectively with children. In order to control for as much measurement error as possible, content validation procedures should be carried out before moving on to other types of validation procedures. The purpose of this study is to take the first steps in the development of such an instrument by gaining a better understanding of how children describe and experience optimal challenge. From there, items for a self-report instrument can be generated by using children's words and descriptions to better reflect the construct of optimal challenge from a child's perspective and thus enhance the content validity of the instrument.

Method

To gain an understanding of how children described physical activity experiences as they pertained to their perceptions of optimal challenge, qualitative methods were primarily used. The data collection process used interview techniques and various sources of triangulation to demonstrate the trustworthiness of the data.

Participants

Participants were recruited from Grades 2 (n = 16) and 3 (n = 15) at a local elementary school located in an urban setting. Prior to the study, children from both grades (n = 31) were given a cover letter describing the study and an informed consent form for the parents and child to sign if they agreed to participate. A total of 28 informed consent forms were returned. One participant decided to withdraw from the study leaving a total of 27 (15 M; 12 F) participants. Fourteen were from Grade 2 and 13 were from Grade 3.

Methods of Data Collection

A number of methods of collecting data were used in this study. Field notes were used to record ideas, strategies, reflections, hunches and to note patterns that emerged. Field notes were completed as soon as possible following each physical education class and after each interview. These notes were comprised of descriptive and reflective comments to keep track of how the project was developing, how the research plan had been affected by the data, and to reflect on how the researcher had been influenced by the data (Bodgdan & Biklen, 1982).

An interview guide was developed based upon current motivational theories that integrate optimal challenge. The interviews were used to gain a better understanding of the participants' subjective experiences pertaining to their motivation during the selected activities in physical education classes. All interviews took place within 25 hours of participants finishing their physical education class and all interviews were audio-taped and then transcribed verbatim. To stimulate recall and to provide concrete examples of their physical education experiences, videotaped segments of the participants' preceding physical education class were shown to them during the interview. The interviews were conducted by the researcher who was also the children's physical education teacher.

Although the investigator has had previous experience interviewing children who he has taught (Mandigo & Couture, 1996), there still existed the potential problem of social desirability response bias during interviews. Participants may have felt they had to provide the answer that the interviewer wanted to hear rather than their true feelings. Every attempt was made to assure the children that there were no wrong answers and that the interview was not a test. Despite the potential effect of socially desirable responses, it is believed that having established a positive rapport with the children served as an advantage to soliciting honest answers. Children were encouraged to answer the questions honestly and were assured that their answers would be held in strict confidence.

Structured interviews with room for probing questions were used to gain a better understanding of children's physical activity experiences and how they related to their perceptions of optimal challenge and perceptions of non-optimally challenging tasks. Questions relating to children's descriptions of easy tasks, just right tasks, too hard tasks and tasks exhibiting various relationship combinations of perceived skill and challenged were asked (see Figure 2.1).

To help facilitate discussion during the interviews, participants were asked to rate on a 5-point Likert scale their perceived skill level, perceived challenge of the activity, and the degree to which they thought the activity was just right, too easy, or too hard for them in a journal after each class. This instrument was designed around Csikszentmihalyi and Larson's (1987) Experience Sampling Form (ESF) which measures perceived skill and challenge and Harter's (1978b) measure of perceived difficulty. The items in this "journal" were not scored but rather, they were used to stimulate conversation during the interviews and to allow participants to write down or draw comments about their participation in the various activities in order to help them remember how they felt during the activity.

Procedures

A pilot study was conducted with five participants from a children's physical activity program prior to the study. The purpose of the pilot study was to test out the method, to practice filming and interview techniques, to help develop the wording for the journal, to practice doing a content analysis, and to help finalise interview questions that were used in this study.

At the conclusion of the pilot study, the main study commenced. For each class taught as part of the study, participants completed one page from their journals immediately following participation in a pre-selected activity. Participants were encouraged to ask questions if they did not understand and were encouraged to work alone. Participants could also write or draw comments about the activity on the back of the page.

To promote interest in a variety of areas, activities from five different activity dimensions (i.e., gymnastics, games, individual activities, alternative environment, dance) from the Alberta Physical Education Curriculum (Alberta Learning, 2000) were used. The activities ranged from discrete tasks (e.g., drill-like activities that are skill based) of various difficulty levels to more open-ended activities (e.g., games, gymnastic routines, created activities). A list and description of these activities are provided in Figure 2.2. Because some children started the study later than others (e.g., brought consent forms back late, absent from class), there are more than 10 activities listed. Participants only completed a journal entry for the first 10 activities in which they participated.

Within approximately 25 hours of participating in the pre-determined activity, one student from each grade was individually interviewed. Upon entering the interview room, participants were told of the purpose of the interview, their rights as a participant in the study, and were asked permission to audio-tape the interview. Toy animals were placed throughout the interview room to make the children feel more comfortable with an unfamiliar environment. As well, the interviewer sat in a chair that allowed him to be at eye level with the participant. Participants then watched a brief video-clip of them participating in their preceding physical education class. This took anywhere from 5 to 10 minutes. After watching the video, they were asked if they would like to watch it again for clarification. Participants were then interviewed with respect to five major areas: 1) their description of what they were doing during the activity on the video; 2) a description of how they felt during the activity on the video; 3) their descriptions of whether the activity was too hard, too easy, just right, challenging, whether they had enough skill to do the activity and whether how they felt their skills matched the challenge of the activity; 4) their understanding and subjective interpretation of those words; and, 5) descriptions of other activities that match the descriptions of just right, too easy, too hard, challenging, and had enough skill. All interviews lasted for a duration of approximately 20 to 45 minutes.

Field notes on interviews and classroom activities were also collected to assist the researcher in evaluating how the study was going. As a result, several new probing questions were included to help explore the emerging categories that were developing throughout the study. This type of approach of starting one's analysis in conjunction with data collection is a strategy encouraged by several qualitative investigators (e.g., Bogdan & Biklen, 1982; Creswell, 1994).

The researcher was both the lead investigator and one of two instructors who "team-taught" the participants' physical education class. Considerable debate surrounds the extent to which qualitative researchers should be involved with the participants. Some argue that it is important to get close to the participants and become a major part of the environment (e.g., Jackson, 1995). Others caution that becoming too close to the participants may bias responses provided by participants (e.g., Creswell, 1994). The position adopted for this study supports the former view. This was done for several reasons. One of the most important components of conducting a good interview with children is to establish positive rapport and trust (Garbarino & Stott, 1989). By being close to the participants, an investigator gains a unique perspective of the environment and the characteristics of the participants in the study that can help enrich the data collection and analysis. The researcher in this study team-taught the physical education classes with another instructor four months prior to and during the course of the study. Although there are several advantages and disadvantages to this design, it was felt that with children, it is important to establish a close rapport and trust prior to conducting one-to-one interviews. It was also felt that the rapport the researcher could establish with the students would enable him to ask probing questions that were more reflective of the abilities and experiences of the participants. In addition, by being ingrained into the environment, the researcher was better able to ask pertinent questions surrounding activities that took place during class. The disadvantage to this approach lies in the potential for the researcher to "indoctrinate" participants about their perceptions of optimal challenge. Because the researcher believes in the potential benefits of optimally challenging children during physical activities, he may have subconsciously "forced" his beliefs of optimal challenge on the children themselves. Although this is a potential weakness of this study, the interview questions were focused around the participants' perceptions of the phenomena of optimal challenge based upon real-life experiences and their interpretation of the concepts that make up optimal challenge experiences.

To control for the potential confounding nature of having the researcher being so close to the participants, five children from a different school where the researcher was not the instructor were interviewed using the same procedures as those indicated above except for the use of the self-report journals. A description of the type of activities participants took part in is listed in Figure 2.2. This "replication" study took place after all data had been analysed from the initial study. This type of data source triangulation adds to the trustworthiness of the results.

Data Analysis

The purpose of qualitative research is to "... objectively report the perceptions of each participant in the setting" (Morse & Field, 1995, p. 142). As such, the data analysis that is used must reflect as accurately as possible the "voices" of the participants. For the purpose of this investigation, content analysis procedures as suggested by a number of authors (e.g., Bogdan & Biklen, 1982; Creswell, 1998; Jackson, 1992; 1996; Morse & Field, 1995; Patton, 1990) were used to retain as much as possible, children's choice of words as they pertain to the construct of optimal challenge. Each interview was transcribed verbatim. Once all interviews had been transcribed, all of the transcripts and field notes were read through once. During this process, notes were taken pertaining to potential higher-ordered coding categories. Numbers were then assigned to these preliminary coding categories and then assigned to units of data during the second time the data were read. An inclusion statement was written which indicated the criteria for including a statement in a particular category. This inclusion statement was important to ensure similar pieces of data were grouped together and dissimilar pieces of data were placed into other categories. This type of approach is consistent with the comparative pattern analysis (Patton, 1990).

In order to gain a better understanding of children's interpretations of optimal challenge, the higher-ordered categories reflected the structured questions used in the interview. Then, statements coded within a particular higher-order category were cut out and grouped together into separate file folders using the Cut-Up-And-Put-In-Folders Approach (Bogdan & Biklen, 1982). The categories that were assigned were descriptive labels that were intended to identify the content of the interviews (Morse & Field, 1995). Statements within these files were then content analysed into sub-categories and subsequent lower-order categories using a similar procedure as that used to develop the higher-order categories. Several copies of the interview transcripts were made because some of the quotes could fit into more than one category (i.e., either higher-order, sub-category, or lower-order category). Using the interview questions to formulate the higher-ordered categories and then content analysing responses into lower-ordered categories is commonly referred to as a question analysis (Morse & Field, 1995). Thus, the lower-ordered categories that emerged reflected the most inductive portion of the data analysis in that they represented participants' descriptions and responses to the more general higher-ordered categories. Once all codes were assigned, all coded data were then read through and revisited several times to ensure the initial codes accurately reflected the interview content. Revisions were then made where necessary following this process and after receiving feedback from a peer reviewer who reviewed the original data analysis and subsequent revisions.

Results

As Figure 2.4 demonstrates, a total of 8 higher-ordered categories pertinent to the construct of optimal challenge emerged from the data. These higher-ordered categories reflect the structured questions asked during the interview. These categories were further divided into three sub-categories which provided descriptions of the category (see Figure 2.5). outcomes as a result of participating in the activities under that category (see Figure 2.6). and participants' preferences for activities which fell under the categories (see Figure 2.7). These sub-categories reflected the type of probing questions that were asked during the interview (e.g., description: what does the word challenge mean to you?; outcome: how did you feel when it was too hard?; preference: did you prefer it when it was easy?). Subsequent lower-ordered categories that emerged from within the sub-categories reflected the most inductive component of the data analysis in that they were intended to represent as close as possible the comments provided by the participants. For example, how did participants describe challenging activities or what was the outcome when an activity was too hard? As a result, some lower-ordered categories emerged in more than one higher-ordered category. For example, some participants described activities where skill equals challenge as just right while others used just right to describe challenging activities. It was felt that this was the best way to analyse the data so that a clear indication of how children describe, experience outcomes, and prefer activities pertaining to terms commonly used to make up the construct of optimal challenge could be obtained and compared to existing frameworks and terminology. The number of times a lower-ordered category was included is indicated in parentheses in these summary tables. The following is a discussion of each higher-ordered category with the sub-categories of description, outcomes, and preferences. As well, thick descriptions of the various categories are provided to better reflect the type of language children used. This type of approach has been suggested to help increase the trustworthiness of the interpretations (Patton, 1990).

Just Right

It has been suggested that when an activity is optimally challenging, it is "just right" for the participant (Csikszentmihalyi, 1975). Thus, questions surrounding children's interpretation for this term were asked. To be included in this higher-order category, participants had to either mention or describe an activity that was just right for them or provide a response to questions relating to just right physical activities.

Description. When asked to describe a physical activity that was just right for them, the most common response by participants was that such activities were not too easy and not too hard. For example, when asked whether an activity was too hard, too easy or just right for him, a participant replied: "I would say it was just right. It was kind of hard and kind of easy, so it just suppose in the middle [*sic*]". As well, when describing why basketball was just right, this participant said: "... it's challenging, but it's not too hard or too easy for you".

The idea that just right activities are perfect was indicated by several participants. Children often indicated that they would substitute the expression "perfect for me" to describe activities that were just right. For example, when describing a climbing wall that was just right, this participant said:

> "What you look for in a wall is you don't ... just look for an easy. What you are mainly looking for is where there is some hard... grasping places that you have to grab onto and some easy that you can grab onto. That means when I look at that and see that wall, that gives me perfect [*sic*]".

As well, participants indicated that a just right activity was one that was "good for them". Many indicated that such activities were good for their skill level. One individual used the analogy from the Goldie Locks and the Three Bears nursery rhyme to indicate that just right activities were just like the porridge that was not too hot and not too cold.

Another common description of just right activities was challenge. For example, one participant indicated that rock climbing was an activity that was just right because: "Well, actually, there are some that are pretty easy, but usually, they are pretty challenging". As well, when asked when an activity is just right, this participant said: "Well, I am good at it, it challenges me, and I like challenge".

For some, however, just right activities were easy for them to do. For example, when asked how he would describe the word just right to his sister, this participant indicated: "Just right would mean, it's easy enough ...". As well, when one of the participants was asked how one could find out if something was just right: "You would say, is this like easy, really easy for you?".

The influence of previous experience in doing a particular activity also contributed to descriptions of just right. For example, when asked why rock climbing was just right, this participant said: "I've had the skills since I was about two".

Outcomes. When participants took part in just right activities, they reported positive outcomes. One of the most commonly mentioned outcomes was experiencing some level of success. The amount of success participants reported varied from participant to participant. For many, being able to do a task well was their measure of success during just right activities. For example, when an activity was just right, others who are watching would know it is just right because: "I was maybe ... doing it well. Like I was supposed to. Like I was going to in the plan". For others, their level of success was gauged by extrinsic factors such as trophies. For example, one girl who described soccer as her just right activity indicated that she knew she was doing well in soccer because: "I keep getting trophies". There were also those who indicated that their level of success during just right activities was just being able to do the activity. For example, when describing ice skating as a just right activity, one girl replied: "You could see that maybe I wasn't falling over too much. or never. never falling over".

Another positive outcome of just right activities was the positive affect that resulted from participating in them. Most participants said that when they took part in just right activities, they felt good. For example, when asked how someone would know if an activity was just right for him, one boy said: "You'd see me really happy playing it and I'd be, I'd have a really happy look on my face".

In addition to positive affect, many participants also indicated that when an activity was just right, they were having fun. When describing what he is thinking about when playing soccer (an activity he considers just right for him), a participant responded: "This is really fun. I'd like to do this all the time and I don't want to go home right now".

Another positive outcome of just right activities was the confidence that participants gained from their participation. One participant who considered taking part in a gymnastics-type activity that was just right for him indicated that if others were watching, they might see "... me doing it really well. You'd see me really confident and it wouldn't seem really, like scared".

<u>Preference.</u> Participants indicated that they preferred activities that were just right and if given the choice, they would keep doing them the same way. Some participants felt that if the just right activity was changed, it would either become too easy or too hard. For example, when asked why a just right activity should not be changed, the response given was: "That's because if it is too easy, it's boring. If it is too hard, um, you sometimes give up or you just walk away and forget it and you can't even play".

Some children would like to see the activity become more challenging. For example, when asked about changing a game that he created that he felt was just right for him, one of the participants indicated that he would add more (bowling) pins to the game to "make it more challenging". For others, they would like to see the just right activity made a little bit easier to increase their chances of success. For example, when describing a gymnastics activity on large exercise balls, one participant indicated that to increase her chances of success, she would like to make it so the balls could not roll away during her routines.

Skill/ Challenge Balance

The most common definition of optimal challenge refers to the balance between skill and challenge (e.g., Ryan & Deci, 2000; Weiss & Bressan, 1985). Thus, information on how children describe, experience, and prefer such activities were sought to better understand optimal challenge from a child's perspective. To be included in this higher-ordered category, participants' responses had to reflect conditions where skills and challenge were perceived to be balanced.

Description. When asked about activities where they felt their skills were balanced with the challenge of the activity, the majority of participants described them as just right for them. For example, when one boy was asked to describe how he knew his skills and the challenge of the activity were equal, he replied: "Because it makes it just right for me because I have challenge and I have skill".

Participants also frequently described such activities as not too easy, not too hard. For example, one participant described activities where skill and challenge were balanced by saying: "It is not too easy and not too hard and you can actually do it once or twice". Some, however, chose to describe such activities as easy while others choose to describe activities where skill equalled challenge as hard.

Outcomes. For many, participating in activities where they felt their skills and challenge were equal produced several positive outcomes. Positive affect was one of the most commonly cited outcomes of participating in such activities. Many indicated that when their skill was balanced with the challenge of the activity, they felt good and that they were having fun doing it. For example, a participant answered that activities where the skills are balanced with the challenge: "... feel like I [don't] want to quit. And I have enough skills at it and it is still fun". When asked why it makes her feel good when participating in an activity where she felt her skills and the challenge were balanced, a participant replied: "Cause you can do it but it's still a little bit hard for you but you can do it". Not only did participants feel good because they could do something challenging, but they also felt good because they felt they were experiencing some level of success. For example, when asked why he prefers activities where his skills are equal to the challenge of the activity, a participant responded: "I'll be able to do it and I won't have to wait a long time [to be able to do it]". Another boy also indicated that "... once I think they [i.e., skill and challenge] are equal then I just try it and it works out".

For some, participating in skill/ challenge balance activities was a source of motivation to get better. For example, when asked how it makes her feel when her skill matches the challenge of the activity, one participant replied: "That I should do it more often ... so I can get way better at it maybe". Another participant also described a climbing experience this way: "The climbing is just a little bit challenging for my skill level, but as I keep trying, this is going to raise, my skill is going to raise higher so they are both equal".

Preference. Because of the positive outcomes that the children associated with activities where their skill and challenge were equal, most indicated that they preferred such activities over ones where their skill was greater or less than the challenge of the activity and would continue to do them. When an activity was not equal in terms of skill and challenge, participants indicated that they would practice to get better or change the activity somehow so that the challenge equalled their skill. For example, to try and make skill and challenge equal this participant suggested: "I try to make it easier, like if I were throwing a ball I would go more close to it and to make it harder, I would go farther away". Others said that when their skills are lower than the challenge, they would practice in order to get them to be equal to the challenge of the activity. An interesting comment by one participant suggested that she does not always like skills and challenge to be equal. When asked when she prefers balanced activities, she replied: "Well I like them a lot. Sometimes I like challenging and sometimes I like easy".

Challenge

The word "challenge" is embedded within any explanation of the optimal challenge construct. Thus, a clear indication of children's interpretations and experiences relating to this word was included as part of the interview guide. To be included in this higher-ordered category, participants had to use the word challenge and provide subsequent descriptions or respond to questions pertaining their understanding of challenge or descriptions of challenging activities.

Description. Some children described challenge as something that is hard. For example, when asked what the word challenge meant to him in the context of rock climbing, a participant answered:

> "Hmmm... it means something's like really hard, and really hard, and so, it like gives you, it's hard to do. Like, climbing up the wall was a big challenge for me because it was really hard to do. If you don't get a good enough grip, you are going to fall. You know how much your body weighs so you got to get a good hold and if you don't get a good hold, then you go, so that maybe what a challenge mean [*sic*]".

For some, when they described something challenging as hard, they referred to it in a good way. For example, when asked what the word challenge meant to him, one boy replied: "It means it is sort of hard and I can do it and I won't be bored doing it". Another participant alluded to the positive feature of hard and challenge by saying: "Like a good challenge is like when you are really trying to learn something but it is hard, or how to do it but you want to know how to do it". For others, however, hard in the context of challenge was viewed negatively. For example, when asked to describe the word challenge, a participant responded: "It means, it is a bit too difficult from all the skills that I have". Another participant also indicated that challenge meant: "It is hard, you can't do it that much. But when you get older, sometimes you can do it".

Challenge was also described by some as something that is not too easy but not too hard. For example, when asked how she would describe the word challenge to her younger brother, one participant replied: "I would tell him challenging means that it is hard and it is easy. Like in the middle for you actually".

The influence of past experience also emerged in the description of the word challenge. For example, when asked to describe what challenge meant, a participant replied: "I'd never done it before, that's what I think challenging means. I never done it before and I don't know how to do it and I need someone to teach me how to do it. That's what it means to me". For most, however, this lack of experience was not viewed negatively. In fact, it was often viewed as a source of motivation to improve. For example, when asked how he would describe challenge to his younger sister, a participant said: "It means, you are not prepared for it and you never, ever tried it before. But after you are done, it's really, really good".

In addition to "hard" providing negative descriptions, some described challenge as something that is not good for you or something that is dangerous. For example, one participant described a bad challenge as: "… something that someone dares you to do like a dog dare or a triple dog dare. They're bad and they're a challenge and they are deadly and they can hurt you". An interesting comment by one individual described a bad challenge as something that was: "really easy for you".

Outcomes. During the course of the study, it started to become apparent that some children viewed challenge as something positive while others viewed challenge as something negative. The idea that something could be a good challenge or a bad challenge soon emerged. Based on the various descriptions of challenge, it is not surprising that participants indicated both positive and negative outcomes associated with challenging activities.

From the positive perspective, some participants indicated that when they took part in challenging activities which were a "good challenge", they felt good about themselves because of the positive affect that it produced. When asked how it would make her feel to participate in an activity that was a good challenge, a participant replied: "It makes me feel like I am going to have a good time and makes me feel like I am fine and it looks fine". Other participants also indicated that when taking part in challenging or good challenge activities, that they had fun because it was exciting and not boring. As one boy described: "If games are challenging for me, it would be a lot funner because if they are easy, they are boring.... [B]oring is too easy for me. Easy is boring, Challenging. That's fun"!

One of the reasons why some participants viewed challenge in a positive light was because of the success they felt after taking part in them. For example, one participant described the outcome of taking part in a good challenge this way: " A good challenge is not like it is impossible, you can't do it. But a good challenge is where you like, you keep on trying and trying and getting it". Others also indicated that when taking part in challenging activities or "good challenges" that they felt it was enhancing their personal development because they were learning something that was good for them and they were getting better at an activity. One of the boys said that: "[I]t makes me feel challenged and it's fun and it makes me feel good because I am getting better at something". As well, one of the participants described challenge as "... something that is harder than most easy things. And it is pretty fun if you think challenges are fun. And it is fairly, fairly, good for you. A challenge ... [B]ecause it teaches you new things". Taking part in challenging activities also made some participants try harder during activities. For example, when asked what she was thinking about when taking part in a challenging activity, a participant replied: "Well, this is pretty hard, but I think I can do it and I will try it". This statement also illustrates that as a result of trying and experiencing success at challenging or good challenge activities, some children felt more confident in their abilities. Finally, being challenged also enhanced the level of engagement within the activity. For example, one participant replied: "Well. I just look at it and then it looks challenging and then I look at it and see that it looks really, really challenging and then I will try it and then if I fail I try again. I never give up. I keep trying until I get it".

Not everyone described the outcome of challenge in positive terms. When describing the outcome of a challenging or bad challenge activity, some indicated their lack of success to do the activity. For example, one participant described an activity that she did during her physical education class as a bad challenge because "... there was too many challenges and the ball was making everyone. like. I didn't know what to cheer for because I didn't know how to play good". Another participant described a bad challenge as resulting in "... getting totally frustrated, you can't do it perfect and you can't do it good and you get really mad". As a result of this lack of success, some indicated that when taking part in challenging or bad challenge activities that it resulted in negative affect such as feeling sad, nervous, or frustrated. Others also indicated that a bad or negative challenge resulted in them feeling contained or not involved in the activity.

Preference. There were those in the study who liked to do challenging things. Many felt that when something was challenging that they would reap the many positive benefits previously described. One of the most convincing arguments was provided by a boy who described how he would choose a wall that he wanted to rock climb.

> "If it looks too easy I walk around to see if there is anything harder cause if there is nothing else that is really harder, I will just try an easy thing and I will just try doing it like faster. Just keep trying doing it, trying to make it challenging because that's like my favourite thing".

As well, some children indicated that they liked challenging activities because they were not boring to do. For example, when asked why she likes challenging things, a participant stated: "Because you can actually challenge yourself, it won't be boring. You wouldn't know if you got it right all of the time".

Some, however, would prefer to either not take part in challenging activities or to change the activity so they could improve their chances of success. For example, after describing football as an activity that was challenging for him, one boy indicated that he would make the ball a little smaller so it would fit into his hand so he could throw it. Others indicated that they would prefer to make the task easier because they do not like challenges. For some, however, they would choose to quit the activity all together. For example, when asked how a challenging activity makes her feel, one participant responded: "Makes me feel like it is hard and I don't want to do it".

Skill

How children perceive their skill level is also an important element in children's perceptions of optimal challenge (Csikszentmihalyi, 1975). Quotes were included in this higher-ordered category if participants responded to questions about skill or made some reference to skill or ability.

Description. When describing the word skill, many participants indicated that they thought it referred to their ability to do something. For example, when asked to describe what skill meant to him, one participant replied: "Skill, well, it means you have like really good and you are really confident and you are going to do it and you really know that you can do it". Another participant also indicated that: "... if M. asked one of us what does skill mean, I would say like you have enough skill to do something. I have enough skill to mountain climb. If you have enough of something to do something. Skill is if you can do something". Others also described skill as something where they had some sort of previous experience. For example, when asked how she knew she had enough skill to do an activity, a participant responded: "I do other things that are the same like that so I know how to do it".

Outcomes. When participants had enough skills, many positive benefits were described. The most common was the sense of success they felt about being able to do the activity. For example, when asked how he knew he had enough skills in soccer, a boy replied: "Cause I am good. I'm pretty good at it. I get goals, I always get the ball, people pass to me". Another participant also referred to experiencing success as an indication of having enough skill in gymnastics: "[I know I have enough skill in gymnastics] because I usually get it right and I don't get frustrated". When participants did have enough skill, they often indicated that it made them feel good and produced some sort of positive affect. Having enough skill led some to believe that they were better than other children. For example, one participant described a situation where he knew he had enough skill if the other kids had a skill level of two and he had a skill level of four.

When participants felt they did not have enough skill, some said this produced positive outcomes while for others, they indicated negative consequences. Those who mentioned positive outcomes associated with a lack of skill said that it was a motivating factor for them to improve. For example, one participant described a gymnastic routine that involved some very complex skills such as backflip, back saulto, and cartwheels. He indicated that even though he did not have enough skill yet to do those actions, he would keep practising to try and get better. Another individual who described how he learned to stop on ice skates indicated at first he could not do it so he started off very slowly and then gradually increased his speed as he became better at stopping.

Not having enough skill led to negative outcomes for some. For example, one participant indicated that not having enough skills made him feel like a "dork" and wanting to try a different sport that had the right skill level for him. Another participant, indicated that he might try the task for a while, but if he does not start to experience some sort of success, then he would likely quit. This lack of success was also indicated by other participants in the sense that without enough skill, they could not do an activity and it was hard to do.

Preference. Most participants indicated that they preferred activities where they had enough skill. When participants described activities where they did not have enough skill, they indicated that they would likely not participate any longer. For example, when asked what he would do if he took part in an activity where he did not have enough skill, a participant replied: " Um... like, go to a different sport or go to a different place and see if they've got the right skill level for me".

Challenge Greater Than Skill

Unbalanced activities where a participant perceives the challenge of the activity to be higher than their skill level has been suggested to cause anxiety in individuals (Csikszentmihalyi, 1975). Thus, it was felt that in order to obtain a clear understanding of the construct of optimal challenge, description, outcomes and preferences for non-optimally challenging activities were needed.

Description. As might be expected from pervious research, when participants felt the challenge of the activity was higher than their skill level, they often indicated that such activities were too hard for them to do. For example, one participant indicated that when the challenge of the activity was higher than her skill to do the activity that: "It means that it is mostly too hard and I can't do that very well ... I have no idea how to do this and this is boring because I don't know what the rules are". Other participants indicated that when the challenge of the activity is higher than their skill that they did not have enough skills to do the activity. For example, one participant described such an activity as: "...you have, the sport that you are playing challenges you a lot but you don't have enough skill for it".

Outcomes. For many, the outcome of such activities was negative. The lack of success associated with challenge being higher than skill was one of the most common consequences. For example, one participant suggested that in such activities: "It will be too hard and you might not be able to do it in a long time". For others, they simply indicated that it was something they just could not do because they were not good at it. As a result, some participants indicated that challenge being higher than skill resulted in negative affect such as feeling "yucky" and "not very, very, very, very good". Others also indicated that participating in such activities would make them: "... want to quit before you got into a bad area" and would be boring because they would have no idea how to do the activity.

For some, however, the outcome of challenge being higher than skill motivated them to keep trying to get better. For example, one participant described certain types of climbing walls as an example of the challenge being higher than his skill level. When he sees such a wall he thinks: "It looks too hard for my skill level, but if I fail, then I will just keep trying, and trying, and trying. And as you keep doing it, your skill keeps getting higher and higher and higher until it just are the same".

Preference. Despite some of the positive outcomes, most indicated they would change an activity where the challenge was higher than their skill. For example, one participant indicated that he would lower the basketball nets so that the challenge would equal his skill level. Others indicated that they would like to practice their skills in order to improve and meet the challenge of the activity.

Too Hard

Non-optimally challenging activities have also been referred to as being hard to do (Reeve, 1996). When the challenge of the activity is higher than what the participant perceives he/she can do, it is often characterised as being too hard. Hence, participants' descriptions of activities that are too hard for them were sought to shed light on how participants described such activities.

Description. The most common description of activities which were too hard reflected concerns over the equipment that was being used. In basketball, for example, many felt the net was too high for them to score a basket. Another participant also described a climbing wall that was too hard for him because the holds on the wall were small and difficult to hang onto. Too hard was also described by some as something that is challenging or difficult for them to do. Things that were too hard were also described with respect to the lack of experience they had with such activities. For example, one participant indicated that he used to think ice skating was too hard because: "I started when I was four and that would mean like I'm a kid who barely knows how to talk and move and walk, balanced on, on bladed skates on a frictionless floor". Another participant indicated that when the teams are unfair, it becomes too hard. For example, in one game, a team ended up playing three against five and one of the participants on the team with fewer players felt this was too hard and unfair. Other descriptions of too hard included not knowing how to do an activity and risky.

Outcomes. When an activity was too hard, many participants indicated that they felt they could not experience success at the task. For example, when describing tennis as an activity that was too hard for him, one participant indicated that it was hard because: "I don't know how to get it over the net very good or how to get it to my partner very good". As a result of this lack of success due to the activity being too hard, many participants expressed concern over the negative affect that they felt. For example, when asked how she feels inside when she is ice skating, one participant answered: "Like grrr, I can't do it". Other types of negative affect included feeling frustrated, getting angry, not being very fun, sad, embarrassed, silly, helpless, and anxious.

There were some, however, who were motivated by activities that were too hard. For these individuals, they felt that they would likely keep trying an activity that was too hard and practice in order to get better. Another participant indicated that an activity that was too hard made her feel weird, but in a good way.

Preference. Despite some positive outcomes of too hard and some who suggested that they would not change an activity that was too hard for them, the majority of participants reported that they would change such activities in order to improve their chances of success. For example, those participants who felt basketball was too hard for them would lower the height of the net so they could experience success. Others reported that they would try and improve their skills so they would eventually be able to do things that were too hard for them. As well, some indicated that they would change the rules of the game if it was too hard. For example, one participant suggested that in basketball, instead of shooting the ball into the net that was too high, she could just dribble around in different ways.

Some, however, reported that they would only make changes to an activity that was too hard if it was an activity that they liked. One participant suggested that: "Well sometimes hard things are good. Sometimes you want to keep them the same, sometimes you want to change them. When you really like that hard thing and when you really want that challenge". If they did not like the game, however, some would rather quit all together. For example, one participant indicated that miniature golf was an activity that was too hard for him but he would not make any changes to the game. If he was playing British Bulldogs, however, and he was one of the last persons to be caught, he would just quit because it would be too hard to get past all of the "Bulldogs" in the middle.

Skill Greater Than Challenge

On the opposite side of the spectrum, activities can also be unbalanced when participants perceive their skills to be higher than the challenge of the activities. Therefore, to gain a complete picture of the optimal challenge construct, participants were asked about activities that fit this description. To be included in this higher-ordered category, participants had to provide descriptions, share outcomes, or indicate their preferences for such activities.

Description. The most common description of skill being higher than challenge was that the activity would be easy. For example, one person replied that when skill level was higher: "It means it will be way too easy and it won't be a challenge". Some other descriptions included too much skill, skill harder than challenge, challenge is hard to get at, not equal, and just right.

Outcomes. When participants perceived their skill to be higher than challenge, most people felt that they could do it and hence, experience success at the task. For example, one participant provided an analogy from a children's book of how it feels when skill is higher than the challenge: "Like the tortoise in the tortoise and the hare. Like I can do anything". Participants also said that they felt some sort of positive affect by indicating that these activities made them feel good about themselves and that they were fun. However, participating in these type of activities was boring for some and decreased their motivation to continue. After asking one participant how such activities made him feel, he replied: "This is almost too easy and I'll keep on doing it but if it gets too easy, I'll quit … because it would get too easy and yawn, this is too boring".

Preference. In terms of preference for activities where their skills are higher than the challenge, there was a mix of responses from children. Some indicated that they liked these type of activities and would not make any changes because it provided them with a sense of relief from "tricky" activities in which they take part. Another participant also said that she would not change an activity where her skill was higher than the challenge because sometimes, she just liked to do simple things and other times, she liked to do hard things. As well, one individual said that he would prefer to have his skill higher than the challenge so that it could be easier and he could do it more and have fun.

There were others, however, who indicated that they would change the activity somehow to either make the activity harder or equal. These individuals wanted to make the activity harder in order to provide a bit more challenge. Many cited that they would increase the challenge to a point where their skills and the challenge of the activity were equal or where it was half hard and half easy.

Too Easy

Like too hard, too easy is often associated with activities deemed not optimally challenging (Reeve, 1996). Previous research and responses by the participants in this study suggest that the condition of skills being higher than the challenge of the activity is one that is easy to do. To further examine the meaning of too easy, participants were asked to indicate how easy an activity was during their physical education class and to describe physical activities they considered too easy for them to do.

Description. The most common description of too easy referred to activities in which the participants had previously done quite often. For example, when providing a description of an activity that was too easy for her, one participant responded: "I practice it a lot of time and I am good at it". Another participant indicated that too easy is: "... if I were doing the jumping with the rope, that would be too easy, but easy is where you already know how to do it".

Another common description of too easy was not challenging/ not hard. For example, when asked why cartwheels were too easy for him to do, one of the participants replied: "You don't have a challenge with it". When asked to provide another word for too easy, one participant also indicated: "Too easy .... ummm... that the word for too easy is it is not challenging you". Similarly, others described such activities as being simple. For example, when asked why rolling a ball back and forth to a partner is too easy for him, the participant replied: "Cause you just sit there and it's too cinchy. You can catch the ball".

Outcomes. When an activity was too easy there was a mix of positive and negative outcomes suggested by participants. Some felt that such activities would be boring. For example, one participant indicated that an activity that was too easy would make him feel bored because it was not a challenge for him to do. Other participants also indicated that easy activities are not fun to do and resulted in negative affect.

For others, the outcome of activities that were too easy was viewed as fun and produced positive affect such as making them feel happy and good about themselves. An activity that was too easy was one in which participants felt they could experience success quite easily. For example, one participant described soccer as an activity that was easy for her to do. She indicated that playing soccer made her feel good because she can score lots of goals and then her team wins. Another participant also indicated that if people were watching him play baseball, they would know that it would be too easy for him if he were able to make a triple play, one of the hardest plays in baseball.

Preference. As might be expected, because of the wide range of preferences associated with activities that were too easy, some participants indicated that they preferred such activities while others said they would make the activity harder. For example, one participant indicated that she liked easy activities and would not change them for fear that the new, more difficult activity would be too dangerous and she may risk injuring herself. Others, however, expressed a desire to change easy activities. For example, when asked why he would change an activity that was too easy for him, one participant replied: "... when something is so easy for me, it is not right for me because I like challenges". As well, another participant who described skipping as an activity that was easy for her reported that she would make it harder because: "I am already a good skipper and I think I should make it more challenging by doing more tricks, really hard tricks that I do". For some, however, activities that were too easy resulted in a desire to quit the activity all together. For example, when asked what he was thinking about during an activity that was too easy, one participant replied: "That I would want to stop it pretty soon".

Trustworthiness

In addition to providing rich descriptions through the use of quotes, various forms of triangulation as suggested by Patton (1990) were used to demonstrate the trustworthiness of the data and the interpretation of the findings. The forms of trustworthiness used for this study were: analyst triangulation; data source triangulation; and, theory/ perspective triangulation.

Analyst triangulation. Analyst triangulation uses multiple as opposed to singular analysts (Patton, 1990). In the current study, this type of triangulation was established by asking a peer reviewer to code a percentage of the data. The peer reviewer was a fellow graduate student who had an extensive background in qualitative research in the area of sport psychology. Initially, the external checker was given approximately 25% of the statements that were contained in each of the higher-ordered categories and asked to use the sub-category and lower-ordered sub-category codes to code each of the statements. Revisions were made based on the external coder's comments and a consensus was reached on how to best present the data. Once the revisions were made, the external checker was given the higher-ordered categories along with the sub-categories and asked to code 20% of the lower-ordered sub-categories to ensure that appropriate revisions to the most inductive portion of the data analysis were made. An inter-rater agreement rating of 79% between the researcher and external coder was established during this second phase. Minor revisions were made to the coding categories based on comments and suggestions by the peer reviewer.

Data source triangulation. This type of triangulation involves triangulating different data sources (Patton, 1990). For the current study, it was used as a mechanism to ensure that the researcher did not confound the responses from participants he taught and to examine similarities in responses. A total of 5 children (3M; 2F) attending Grade 2 and 3 classes from a different elementary school provided informed consent and took part in a replication of the study. At the end of one of their physical education classes, one to two children were asked if they would be willing to be interviewed about their physical education class and experiences in other sports or physical activities in which they participate. These interviews took place within 3 hours of their participation in physical education. A brief description of the content of their physical education classes is provided in Figure 2.2. The videos of the children participating in the class were used as a source of stimulated recall for the participants. An interview guide identical to that used in Study 1 was then conducted (see Figure 2.8). All interviews were transcribed verbatim and content analysed in a similar fashion to that used in the initial study. The same peer reviewer from the initial study was once again asked to code 20% of the data using the lower-ordered categories that had been generated. An inter-rater agreement of 85% was established. Revisions were made where necessary based upon the peer reviewer's comments.

Out of the 106 lower-ordered categories that emerged from across the 8 higher-ordered categories from the replication study, 74 (or 70%) of them also emerged in the initial study. These categories are indicated with a star (*) in Figures 2.3, 2.4, and 2.5. Given the fact that: "... triangulation of data sources within qualitative methods will seldom lead to a single, totally consistent picture" (Patton, 1990, p. 467), there is consistency in the overall pattern of data from the replication when looking at the categories that emerged most frequently across both studies.

Despite the similarities, slight differences need to be addressed. Because the purpose of the study was to present the data using terms that children use, similar types of words were used in both studies but labelled differently due to inter-participant differences. For example, in the replication study, one of the outcomes that emerged from just right activities was "being proud". Although this category did not emerge in the initial study under this higher-ordered category, similar outcomes such as "being confident" did emerge. Another example was the lower-ordered category of "don't want to do it" that emerged in the initial study as preference for activities that are too easy. In the replication study, a similar category of "want to try something new" emerged as a preference under the same higher-ordered category, but categorised that way to retain the words that the participants used as much as possible.

In other cases, similar lower-ordered categories emerged, however, they emerged under different higher-ordered categories. For example, under the challenge higher-ordered category, "equal teams" emerged as a description in the replication study. This same lower-ordered category also emerged under descriptions of just right in the replication study. Another example is the lower-ordered category of receiving positive feedback from others that emerged as an outcome under the too easy higher-ordered category in the replication study, but emerged under the skill higher-ordered category in the initial study.

Another important difference between the two studies is that in the replication study, participants did not use the word "challenge" to describe Just Right, Too Easy and Too Hard activities. This is cause for concern given that the word "challenge" in conjunction with skill is used to describe the construct of optimal challenge. Perhaps participants in the initial study had been exposed to the word challenge more due to their type of school. These participants attended an elementary school that is located inside a university where a problem-based approach to learning is used. As a result, students are often encouraged by their teachers to challenge themselves and to work at a level that is appropriate for them. As well, the lead researcher was also one of their teachers and as such, more than likely used the word challenge quite a bit without knowing. Thus, it may be cause for concern that some children may not voluntarily use the word challenge. Due to practical concerns to conduct the study and arrangements made between the participating teacher from the replication school and the researcher (i.e., only being able to interview five children in the replication study), saturation may not have been reached.

That is to say, new themes from the replication study emerged and some of the data from the replication study did not match the initial study. However, due to the nature of this study and the type of analysis that was conducted (i.e., lower-ordered categories were kept as close as possible to the vocabulary children used), this is not surprising. Future studies may wish to use more participants and make appropriate arrangements so that more data can be compared and contrasted.

One of the more obvious differences that emerged from the replication study was a participant who expressed very clearly his desire to do easy tasks so that he could look good in front of others to gain their positive approval. This is evident across several higher ordered categories. Although there were those in the initial study who indicated they preferred easier tasks and did not always like hard ones, the participant in the replication study was very clear and adamant about his preferences for easy tasks in order to look good in front of others and to avoid embarrassing himself. For example, many of the outcomes he expressed indicated that an activity was just right when he was able to impress others and make new friends. He went on to suggest that his desire to have enough skill to do an activity is influenced by his desire to look good in front of others. He indicated that he liked easy things because he receives positive feedback from others. For example, after describing soccer as an activity where his skill was higher than the challenge, he indicated that soccer makes him feel good: "Because people will be like, yeah, and saying I am good and all stuff like that". He went on to say that if something is too hard for him, he would be embarrassed in front of others because he could not succeed at the activity. He also indicated that he preferred physical activities where his skills were higher than the challenge because: "... people won't say I was kind of bad at that game and I just want to be good".

Theory triangulation. The final type of triangulation mentioned by Patton (1990) is that of theory triangulation. This type of triangulation involves using different theoretical perspectives to interpret the same data. As mentioned in the introduction portion of this paper. three theories of intrinsic motivation have been used to examine the phenomena of perceived optimal challenge and its impact on intrinsic motivation (see

Figure 2.7). Therefore, these three theoretical perspectives are used in the discussion section to gain a better understanding of the results and future directions of this research.

Discussion

The main purpose of this study was to uncover how children describe experiences which are optimally challenging (i.e., perceptions of skill level are equal to perceptions of the challenge) so that developmentally appropriate items for a self-report instrument could be generated. Based on many of the children's comments, it can be argued that children can describe instances where they have experienced optimal challenge. Although many children did not actually say that they experience "optimal challenge" in those exact words, they did describe many experiences that were theoretically similar to this construct. The evidence for this lies in the descriptions by many participants who said that they preferred activities which were not too easy or too hard. According to several humanistic theories of motivation (e.g., Csikszentmihalyi, 1990; Deci & Ryan, 1985; Harter, 1978a), optimal challenge is the most motivating type of activity. Many children in this study claimed that when the challenge of the activity was higher than their skill level or when an activity was too hard, they would quit. Similar responses were also provided by some who said that they would quit activities where they felt their skill was higher than the challenge of the activity or when it was too easy for them. For both of these conditions, many claimed that they would prefer to change the activity somehow so that it was just right or where their skills matched the challenge of the activity.

However, important individual differences were also uncovered. For example, some participants indicated that they would prefer activities where their skill is higher than the challenge of the activity. Previous research supports the positive subjective experience of such conditions (e.g. Haworth & Evans, 1995; Mandigo & Couture, 1996; Stein, Kimiecik, Daniels, & Jackson, 1995). This finding suggests that there may be situational and dispositional differences in the preference for optimally challenging activities. For some, optimally challenging activities may not produce the highest levels of intrinsic motivation. For example, previous research has suggested that individuals who hold an ego-orientation may prefer easier activities in order to protect their sense of self in front of others (Sarrazin & Famose, 1999). This finding is supported by the clear message by the participant in the replication study who indicated that he would prefer easy tasks so that he could look good in front of others in order to receive their social praise. Reeve (1996) suggested that individuals who prefer to perform in front of others will choose easier tasks to demonstrate their competence as opposed to learners who tend to choose harder tasks to enhance their own personal development. As well, Delignières (1999) suggested that dispositional dimensions such as sex-roles may also influence an individual's sensitivity to optimal challenge.

Situational variables may also impact children's preference for optimally challenging tasks. Harter (1978b; 1992) found that when children were placed into a condition where their performance was controlled by grades, children often chose easier tasks so that they could do the task and receive the higher grades. The idea that rewards can serve to undermine intrinsic motivation is also supported by Ryan and Deci (2000) who put forth a continuum of extrinsic/ intrinsic motivation. Controlling rewards which provide the main source of motivation are closer to the extrinsic end of the continuum while rewards which are intended to inform one about their sense of competence serve to enhance intrinsic motivation. It would seem to reason that if children do not feel constrained by rewards, that when given the choice, they will be more likely to choose more optimally challenging tasks rather than easier tasks where their success is guaranteed.

For the most part, however, participants reported optimally challenging activities as being positive. Support was provided for all three humanistic theories that have included optimal challenge as an important construct for intrinsic motivation due to the enhancement of perceived competence that results from success through such activities. Physical activities in which participants perceived their skills to be balanced with the challenge of the activity were viewed by the majority of children as being the most positive and just right for them. Participants also expressed a desire to continue doing these activities and to modify other activities so that skills and challenges were equal. This finding supports Csikszentmihalyi's (1990) claim that optimally challenging activities are highly motivating. He argued that when participants feel their skill level is balanced with the challenge of the activity, this instils a desire to keep doing the activity and to increase the challenge of the activity as their skill level increases. If an activity is too easy or too hard, children are less likely to persist at the task (Sarrazin & Famose, 1999). As well, if a child continually experiences failure at a difficult task, it can lead to withdrawal. For example, Rogers and Ponish (1987) found that when given choice during a beanbag toss activity, children were more likely to persist at a challenging distance and experience some failure only if they eventually started to experience success. If the child does not start to experience success, it can be assumed that the activity is too hard and modifications need to be made. However, as the results from this study suggest, the amount of tolerance may differ between children. Some indicated that if they did not have enough skill, they would keep trying until they were able to do it. Others, however, indicated they would rather quit the activity than keep trying to improve their skills. This finding combined with previous research provides evidence of the importance of taking into consideration both dispositional and situational variables when examining the construct of perceived optimal challenge.

Many participants indicated that when they did activities that were optimally challenging for them that they felt good about themselves and more confident in their abilities. This finding supports Deci and Ryan's (1985) Cognitive Evaluation Theory which suggests that when individuals experience optimal challenges, their perceived competence is enhanced which in turn enhances their intrinsic motivation to do the activity. Harter (1978a) took this a step further to suggest that optimal challenge in conjunction with success is the key ingredients to enhancing perceived competence and intrinsic motivation. Many participants felt that when the activity was just right or when their skills and the challenge were balanced, that they were much more likely to experience success and feel good about themselves.

Underlying the construct of optimal challenge is the assumption that individuals have a desire to reach what Maslow termed self-actualisation. According to Maslow, the ultimate goal of human motivation is to attain self-actualisation whereby we strive to reach our full potential and to become the people we truly are meant to be (Slife &Williams, 1995). Without this assumption, the construct of optimal challenge would fall apart because people's desire to seek out new challenges once they improve upon their abilities is the underlying feature of optimal challenge. Reeve (1996) suggested that this is due to the desire to seek out activities that will enhance perceived competence. In the current study, many children indicated that they like to be challenged and prefer activities that do challenge them to a point where the activities are not too hard. Children often indicated that they would change an activity to make it more challenging for them or would continue doing an activity as long as it challenged them. However, the results also seemed to suggest that if the challenge got too high, then this may be a source of amotivation. Again, further research is needed to examine the influence of dispositional and situational factors on individuals' desire to reach self-actualisation.

Although many did indicate their desire to be challenged and to make activities more challenging once they got easier, there were those who thought challenge was something that could hinder their performance and enjoyment of an activity. For these individuals, the word challenge meant something hard and something they could not do. As the study progressed, field notes were helpful in recognising this and asking participants if there was such a thing as a good challenge and a bad challenge. For some, there was no distinction, but for many others, there was. While bad challenge referred to things which were too hard to do and which caused negative affect, a good challenge was viewed as something that participants were able to do and which produced positive subjective experiences such as fun. Therefore, it is suggested that when referring to challenge within an optimal challenge framework, that it be referred to as "good challenge" in conjunction with their abilities or skills to do something. As a result of this difference in interpretation of the word challenge, the current operational definition of optimal challenge (e.g., Csikszentmihalyi, 1975) may not be entirely valid for children's research. If we were to simply ask children to rate the challenge of the activity and compare it to how they rate their skills to do an activity, we may get a wide range of responses to similar experiences because of the diverse interpretations of the word challenge. Therefore, operational definitions of optimal challenge need to be re-examined in order to provide a valid reflection of how children perceive the word challenge. Although Csikszentmihalyi's (1975) definition has served many researchers well over the years in understanding individual's subjective experiences at the moment, the operational definition may not carry over well to children's research.

In addition, participants in the replication study did not voluntarily use the word "challenge" to describe their physical activity experiences. Although children in both studies did describe, experience, and prefer challenging activities in similar ways, the children in the replication study had to be asked directly through structured and probing questions. Although this may be a function of the relatively small sample size (n = 5) in the replication study, it may also be cause for concern. Future research may wish to examine whether children do use the word "challenge" voluntarily when describing their physical activity experiences.

The other key variable often associated with optimal challenge is the participants' perception of their skill or ability level. Some felt that when they did not have enough skill, they wanted to quit. Others stated that they would practice to try and get better at an activity. How children perceived their skills did seem important in determining their motivation to continue participation. Some wanted their skill level to be higher than the challenge of the activity while others wanted their skill level to be on par with the challenge of the activity. Therefore, future research is warranted to try and uncover what type of personalities and which type of situations are most conducive to facilitating optimal challenge.

Implications for Future Research

As indicated previously, the purpose of this study was to be able to gain a solid understanding of how children describe, experience, and prefer various physical activity experiences in relation to factors related to optimal challenge. Based upon these findings, items for a self-report instrument can now be developed that use children's' descriptions as a source of content validity evidence (Crocker & Algina, 1986; DeVillis, 1991). Future research may wish to examine more closely, however, whether or not children voluntarily use the word "challenge" when describing their physical activity experiences. The current study could not clearly assess this based upon the fact that the participants in the replication study did not use the word "challenge" unless prompted by the structured interview questions. Despite this limitation, items that reflect as closely as possible children's language can be pulled from the content analyses that were conducted on the initial and the replication studies. By using children's language based on their understanding of optimal challenge, the amount of error variance can be controlled for as best as possible. As a result, score interpretations that result from the self-report instrument will be much closer to reflecting the true score of perceived optimal challenge across the three content areas.

The results also contained valuable information that will help generate hypotheses later on in the construct validation process. As indicated previously, there were suggestions that personal disposition may influence an individual's preference for and sensitivity to optimal challenge. One person might indicate high levels of enjoyment for an easy task while another individual might indicate low levels of enjoyment for the same task. As well, if the same task is conducted in two separate environments that are not identical, potential situational factors may influence an individual's motivation for various levels of optimal challenge. Therefore, future validation research which use concurrent, criterion-related, and divergent validation procedures as sources of evidence need to be sensitive to the current findings when generating hypotheses surrounding test score interpretation for a self-report instrument on children's perceptions of optimal challenge. For example, an individual with a strong ego-orientation and low perceived competence may score low on intrinsic motivation towards optimally challenging tasks. If an investigator is not aware of the individual's motivational orientation, this may confound the test score interpretation that results when correlating scores from a subjective scale for intrinsic motivation or objective measures of optimal challenge to scores for a subjective measure of perceived optimal challenge.

Conclusion

The results from this study form the first steps in creating a developmentally appropriate self-report instrument. By using children's comments, relevant items pertaining to their perceptions of optimal challenge can now be generated to enhance the content validity of the items and control for potential measurement error. The results also provide insight into potential situational and dispositional factors that must be considered later on in the validation process. Had this study not taken place and the items were simply generated based on existing literature, the confidence level that this study has created in being able to generate items that represent the construct optimal challenge as viewed by children would be lacking. By taking the time to listen to the voices of the participants, a much more valid tool can be generated so that research in this area can go forward and better understand the best ways to intrinsically motivate children through children's perceptions of optimal challenge.

Figure Caption

Figure 2.1. Semi-structured interview questions.

A. Building Rapport
1) Introduction

2) Purpose of the interview.

3) Ask permission to:
  a) ask questions
  b) tape-record the interview

B. Video
1) I'm going to show you a video of an activity that you did last class.
2) Watch video
3) Do you need to see it again?

C. Situation Specific Questions
1) Can you describe what you were doing on the video?
  Probe: Prompt for clarification of key points

2) Can you describe how you felt during that activity?
  Probe: Clarification of key points

3) Do you think that activity was:
  i) Too easy for you?
  ii) Too hard for you?
  iii) Just right for you?

  Probe: Describe why you chose that answer and what the term means to them
    How does that make you feel?
    What do you do to: keep it just right OR *to make it less easy or less hard for you?*
    What do the other terms mean and can you think of a physical activity that fits with that description?
    Can you think of other words for: too easy; too hard; just right?
    What are some things I might see when you experience these things?

4) Was the activity challenging?
  Probe: What does the word challenge mean to you?
    How would you describe challenge to someone younger than you?
    Other physical activities that are challenging for you?

5) Did you have enough skill to do the activity?
  Probe: What does the word skill mean to you?
    How would you describe skill to someone younger than you?
    Other physical activities where you have enough skill?

6) Can you think of a physical activity where you have enough skill and the activity is challenging (do a jester with hands of a scale to symbolise that they are balanced)

Probe: Describe the activity
How that makes you feel
A word for this
How you reach this balance
Things I might be able to see when they are equal?

7) How about when they are not equal (i.e., challenge > skill and skill > challenge)?

Probe: Describe the activity
How that makes you feel
A word for this
How you reach this balance
Things I might be able to see when they are equal?

D. Ending

1) Is there anything else you want to say about the activity?
2) Do you have any questions for me?

Thank-you

Figure Caption

Figure 2.2. Descriptions of physical activities for study 1.

Name of Activity	**Activity Dimension**	**Description of Activity**
Running obstacle course	Individual	Set up a running course made up of hurdles and slanted bars
Nuclear Waste	Games	Students had to work together to lift a bucket of balls with ropes that were attached to the bucket and dump them into a barrel.
Tinikling	Dance	Hopping over and onto various pieces of equipment (e.g., mats, skipping ropes. sticks) to music.
Baseball Games	Games	Students created their own game using either throwing or batting skills.
Hoop Aerobics	Individual	The activity had them do various things with hoops to keep them moving with the purpose of increasing their heart rates.
Body Awareness	Gymnastics	Students worked in groups of three and created various shapes (e.g., twisted, curved, straight) with their bodies.
Floor Hockey	Games	Students had to try to knock down one of three pins to score a goal.
Active Living Orienteering	Alternative Activities	Students were given a map of the gym and a score card. They were then to find various locations on the map, go to that station, and do the activity explained on the card at the station when they found it.
Skittles	Games	Series of pins set up in a team's court. The other team then attempts to knock them over by throwing a ball at them.
Create-A-Game	Games	Worked in small groups. Each group had to choose one skill and one shape to create a game that used a maximum of four pieces of equipment.
Gymnic Balls	Dance	Put together a Folk Dance using gymnic balls.
Beanbag Tennis	Games	Using a folded mat as a net, students threw the beanbag back and forth trying to score a point by having it land inside the court.
Create-An-Activity	Games, Dance. Gymnastics	In groups no bigger than 3, students chose 1 skill, 1 shape, and 2 pieces of equipment to create an activity that included all of these on the card. They then were to record the location of their activity number in the appropriate spot on a map of the gym.
Create-An-Activity	Games, Dance, Gymnastics	They set up their activity from last day and had them show and share their activity with another group. Then asked them to combine their activities together to make one big activity.
Give And Go	Games	Played give and go by throwing a ball to a partner who would pass it back and then hit the blackboard which was covered in chalk.

Figure Caption

Figure 2.3. Descriptions of physical activities for replication study.

Name of Activity	**Type of Activity**	**Description of Activity**
Skipping Routine (Day 1)	Individual & Duel Activity	Participants choose 2 skipping actions that they had previously rehearsed and put them into a routine to show the rest of the class
Rope Work (Day 2)	Individual & Duel Activity	Tried out various rope jumping games during the class - worked in small groups

Figure Caption

Figure 2.4. Overall representation of higher-ordered responses and sub-categories.

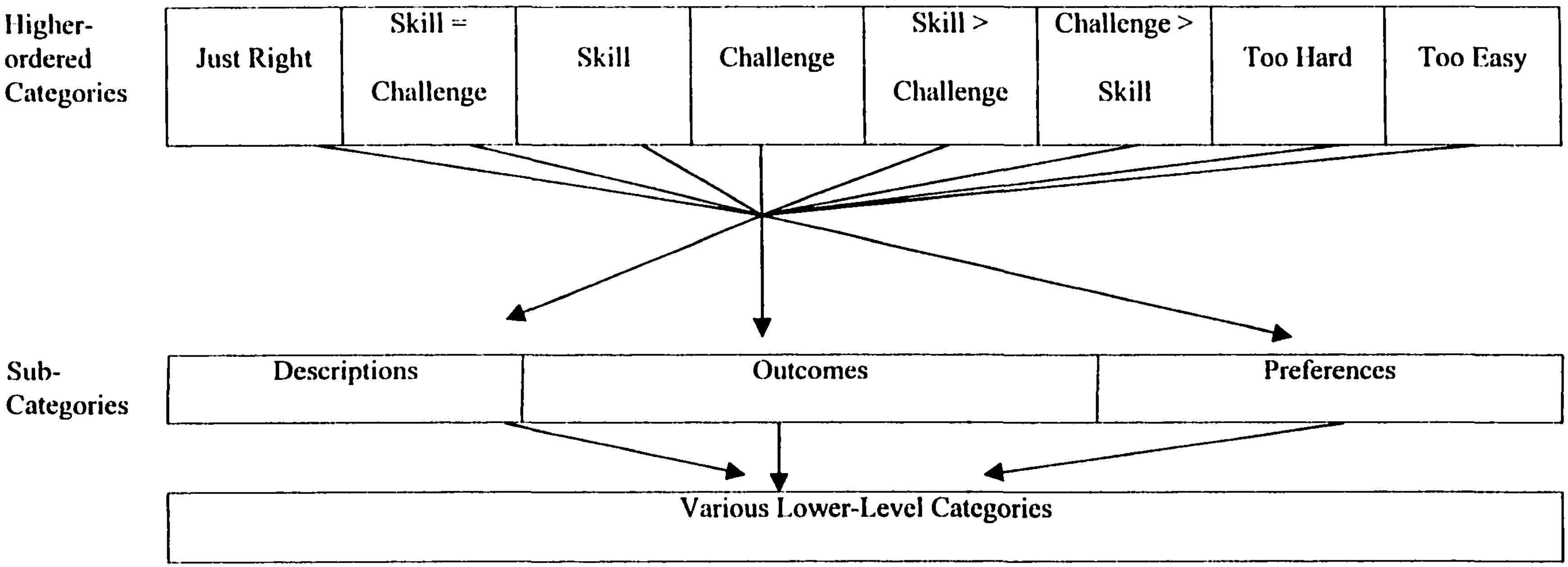

Higher-ordered Categories
Just Right
Skill = Challenge
Skill
Challenge
Skill > Challenge
Challenge > Skill
Too Hard
Too Easy
Sub-Categories
Descriptions
Outcomes
Preferences
Various Lower-Level Categories

Figure Caption

Figure 2.5. Summary of lower-ordered responses under the descriptions sub-category across all higher-ordered categories.

Just Right	Skill = Challenge	Skill	Challenge	Too Easy	Too Hard	Challenge > Skill	Skill > Challenge
Not too hard, Not too easy* (29)	Just Right* (12)	Ability to do something* (12)	Hard* (41)	Prior Experience* (18)	Equipment Concerns* (12)	Hard* (10)	Easy* (17)
Challenging (10)	Not too hard/ not too easy* (7)	Easy* (5)	No Experience* (10)	Not Challenging/ Hard (8)	Challenge/ Difficult (10)	Need more skill* (2)	Not equal (2)
Good for you (7)	Easy (2)	Previous Experience* (4)	Not Hard, Not Too Easy* (10)	Simple* (7)	Lack Of Experience* (6)	Not Equal* (1)	Skill harder than challenge (1)
Easy* (7)	Hard* (1)	Know how to do it* (3)	Risk (5)	Good for you (1)	Don't know how to do it (4)		Too much skill* (1)
Perfect* (5)		Hard (4)	Easy* (3)	Over Average (1)	Unfair Teams* (1)		Just Right (1)
Previous Experience* (5)		Power (2)	Just Right* (3)		Risk (1)		Challenge is hard to get at (1)
Previous Experience (5)			Not good for you (3)				
Good Equipment (3)			Safe to do (1)				
Can make choices (2)							
Better than others (2)							
Even Teams (1)							

Note 1: * indicates that the same lower-ordered category emerged from the replication study.

Note 2: Numbers in parentheses indicate the number of statements that fit into the lower-ordered category.

Figure Caption

Figure 2.6. Summary of lower-ordered responses under the outcomes sub-category across all higher-ordered categories.

Just Right	Skill = Challenge	Skill	Challenge	Too Easy	Too Hard	Challenge > Skill	Skill > Challenge
Experience Success* (38)	Able to do something challenging (28)	Experiencing Success* (31)	Positive Affect* (25)	Experiencing Success Easily* (30)	Lack of success* (34)	Lack of Success* (11)	Experience Success* (10)
Positive Affect* (33)	Positive Affect* (24)	Positive Affect* (15)	Experience Success* (18)	Positive Affect* (14)	Negative Affect* (15)	Moderate Affect (3)	Positive Affect* (8)
Fun* (22)	Experiencing Success* (12)	Want to get better* (4)	Lack of Success* (18)	Boring* (7)	Getting Better (3)	Negative Affect* (2)	Bored (2)
Confident (4)	Fun* (4)	Lack of Success* (3)	Personal Development* (10)	Fun* (6)	Positive Affect (2)	Keep Trying* (2)	Fun (1)
Focused Concentration (2)	Motivation to Get Better (3)	Better than others (3)	Fun* (9)	Better than others (4)		Boring (1)	
	Flow-like experience (1)	Confident (2)	Makes you try hard (7)	Negative Affect (2)			
	Relaxation (1)	Positive feedback from others* (2)	Avoid boredom (4)	Relieved (1)			
	Try My Best* (1)	Trying One's Best (1)	Nervous (4)	Doing what you want (1)			
			Confident (3)	Automatic (1)			
			Getting Better (2)	Feel Energised (1)			
			Negative Affect* (2)	Not Fun (1)			
			Engaged in Activity (2)	Not hard (1)			
			Not boring (1)				
			Not Involved in Activity (1)				
			Exciting (1)				
			Feel Contained (1)				

Note 1: * indicates that the same lower-ordered category emerged from the replication study.

Note 2: Numbers in parentheses indicate the number of statements that fit into the lower-ordered category.

Figure Caption

Figure 2.7. Summary of lower-ordered responses under the preferences sub-category across all higher-ordered categories.

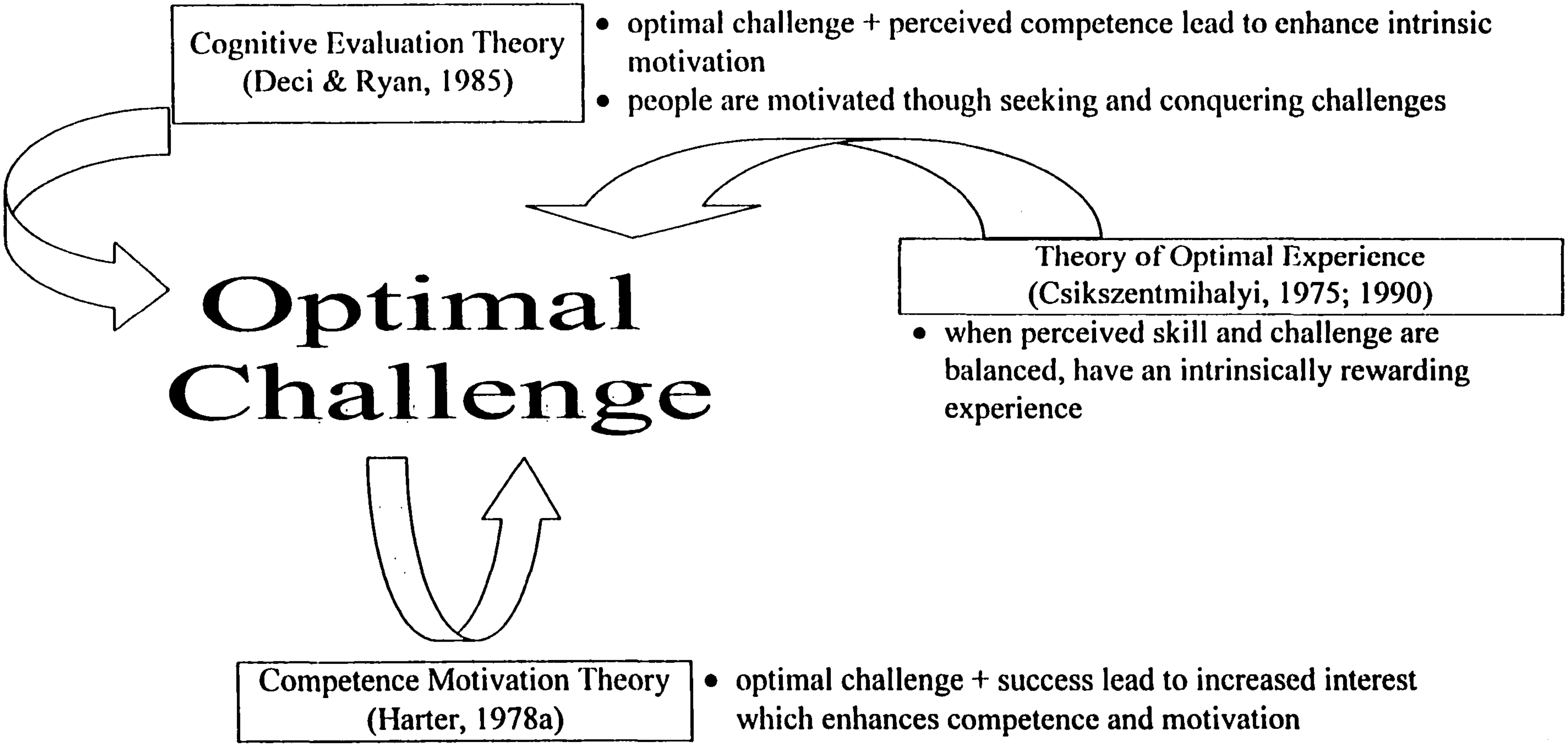
Just Right	Skill = Challenge	Skill	Challenge	Too Easy	Too Hard	Challenge > Skill	Skill > Challenge
Keep it the same* (21)	Would choose balanced activities* (23)	Withdraw if don't have skill (3)	Like challenge* (25)	Make it Harder* (21)	Change it to Improve Chances of Success* (24)	Change it* (8)	Like it* (10)
Make it more challenging (5)	Make a change (3)	Want to do activity if have skill* (2)	Change it to increase success* (7)	Leave it the same* (9)	Have to like activity to continue (3)	Like it* (3)	Would change it* (8)
Change to increase success (2)	Sometimes like challenge/ sometimes like easy (1)		Don't want to do it (3)	Don't want to do it (4)	Want to Withdraw* (3)	Want to Quit (1)	Want to Quit (2)
			Do New things (1)		Keep it the same* (2)		

Note 1: * indicates that the same lower-ordered category emerged from the replication study.

Note 2: Numbers in parentheses indicate the number of statements that fit into the lower-ordered category.

Figure Caption

Figure 2.8. Theoretical triangulation of various intrinsic motivation theories as they pertain to the phenomena of perceived optimal challenge.


Cognitive Evaluation Theory
(Deci & Ryan, 1985)
• optimal challenge + perceived competence lead to enhance intrinsic motivation
• people are motivated though seeking and conquering challenges
Optimal
Challenge
Theory of Optimal Experience
(Csikszentmihalyi, 1975; 1990)
• when perceived skill and challenge are balanced, have an intrinsically rewarding experience
Competence Motivation Theory
(Harter, 1978a)
• optimal challenge + success lead to increased interest which enhances competence and motivation

References

Alberta Learning (2000). Physical education: Guide to implementation. Edmonton, AB: Alberta Learning, Learning and Teaching Resources Branch.

American Educational Research Association. American Psychological Association & National Council on Measurement in Education (1985). Standards for educational and psychological standards. Washington, DC: AERA Publications

Andersen. R. E. (2000). Canadian Medical Association Journal. 163, 1461-1462.

Biddle, S. J. H., & Chatzisarantis, N. (1999). Motivation for a physically active lifestyle through physical education. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand. & R. Seiler (Eds.), Psychology for physical educators (pp. 5-26). Champaign, IL: Human Kinetics.

Bogdan. R. C., & Biklen, S. K. (1982). Qualitative research for education: An introduction to theory and methods. Boston, MA: Allyn & Bacon.

Brustad. R. J. (1998). Developmental considerations in sport and exercise psychology measurement. In J. Duda (Ed.), Advances in sport and exercise psychology measurement (pp. 461 - 470). Morgantown, WV: Fitness Information Technology.

Craig, C. L., Russell, S. J., Cameron, C., & Beaulieu, A. (1999). Foundation for joint action: Reducing physical inactivity. Ottawa, ON: Canadian Fitness and Lifestyles Research Institute.

Creswell, J. W. (1998). Qualitative inquiry and research design. Choosing among five traditions. Thousand Oaks, CA: Sage.

Creswell, J. W. (1994). Research design. Qualitative and quantitative approaches. London, UK: Sage.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Philadelphia, PA: Harcourt Brace Jovanovich College.

Csikszentmihalyi, M. (1975). Beyond boredom and anxiety. San Francisco, CA: Jossey-Bass.

Csikszentmihalyi, M. (1990). Flow. New York: Harper & Row.

Csikszentmihalyi, M. & Larson, R. (1987). Validity and reliability of the experience-sampling method. The Journal of Nervous and Mental Disease, 175, 526 - 536.

Danner, F. W., & Lonky, E. (1981). A cognitive-developmental approach to the effects of rewards on intrinsic motivation. Child Development, 52, 1043 - 1052.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behaviour. New York: Plenum.

Delignières, D. (1999). Perceived difficulty, resources investment, and motor performance. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand, & R. Seiler (Eds.), Psychology for physical educators (pp. 211 - 232). Champaign, IL: Human Kinetics.

Delignières, D., Famose, J., Thépaut-Mathieu, C., & Fleurance, P. (1993). A psychosocial study of difficulty rating in rock climbing. International Journal of Sport Psychology, 24, 404 - 426.

DeVillis, R. F. (1991). Scale development: Theory and application. Newbury Park: Sage.

Dobos, J. A. (1996). Collaborative learning: Effects of student expectations and communication apprehension on student motivation. Communication Education, 45, 118 - 134.

Garbarino, J., & Stott, F. M. (1989). What children can tell us. San Francisco, CA: Jossey-Bass.

Harter, S. (1978a). Effectance motivation reconsidered. Human Development, 21, 34 - 64.

Harter, S. (1978b). Pleasure derived from challenge and the effects of receiving grades on children's difficulty level choices. Child Development, 49, 788 - 799.

Harter, S. (1992). The relationship between perceived competence, affect, and motivational orientation within the classroom: Processes and patterns of change. In A. K. Boggiano, & T. S. Pittman (Eds.), Achievement and motivation: A social-developmental perspective (pp. 77 - 114). Cambridge, UK: Cambridge University Press.

Haworth, J., & Evans, S. (1995). Challenge, skill and positive subjective states in the daily life of a sample of YTS students. Journal of Occupational and Organizational Psychology, 68, 109 - 121.

Jackson, S. (1992). Athletes in flow: A qualitative investigation of flow states in elite figure skaters. Journal of Applied Sport Psychology, 4, 161 - 180.

Jackson, S. A. (1995). The growth of qualitative research in sport psychology. In T. Morris, & J. Summers (Eds.), Sport psychology. Theory, applications and issues (pp. 575 - 591). New York: John Wiley & Sons.

Jackson, S. A. (1996). Toward a conceptual understanding of flow experience in elite athletes. Research Quarterly for Exercise and Sport, 67, 76 - 90.

Mandigo, J. L., & Couture, R. T. (1996). An overview of the components of fun in physical education, organised sports and physical activity programs. Avante, 2, 56 - 72.

Mandigo, J. L., & Thompson, L. P. (1998). Go with their flow: How flow theory can help practitioners to intrinsically motivate children to be physically active. The Physical Educator, 55, 145 - 159.

Messick, S. (1988). The once and future issues of validity: Assessing and meaning and consequences of measurement. In H. Wainer, & H. I. Braun (Eds.), Test validity (pp. 33 - 48). Hillsdale, NJ: Laurence Erlbaum.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed.) (pp. 13 - 103). New York: American Council on Education.

Morse, J.M., & Field, P.A. (1995). Qualitative research methods for health professionals. Thousand Oaks, CA: Sage.

Patton, M.Q. (1990). Qualitative evaluation methods (2nd ed.). London, UK: Sage.

Reeve, J. (1996). Motivating others: Nurturing inner motivational resources. Needham Heights, MA: Allyn & Bacon.

Roberts, G. C. (1992). Motivation in sport and exercise: Conceptual constraints and convergence. In G. C. Roberts (Ed.), Motivation in sport and exercise (pp. 3 - 29). Champaign, IL: Human Kinetics.

Rogers, C. S., & Ponish, K. K. (1987, April). Control of level of challenge: Effects on intrinsic motivation to play. Paper presented at the Society for Research in Child Development, Baltimore, MA.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55, 68 - 78.

Sarrazin, P., & Famose, J. (1999). Children's goals and motivation in physical education. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand, & R. Seiler (Eds.), Psychology for physical educators (pp. 27 - 50). Champaign, IL: Human Kinetics.

Secretary of Health and Human Services and the Secretary of Education, 2000 (2000). Promoting Better Health for Young People Through Physical Activity and Sports. Silver Spring, MD: Centre for Disease Control.

Slife, B. D., & Williams, R. N. (1995). What's behind the research? Discovering hidden assumptions in the behavioural sciences. Thousand Oaks, CA: Sage.

Smoll, F. L., Magill, R. A., & Ash, M. J. (1988). Children in sport (3rd ed.). Champaign, IL: Human Kinetics.

Stein, G. L., Kimiecik, J. C., Daniels, J., & Jackson, S. A. (1995). Psychological antecedents of flow in recreational sport. Personality and Social Psychology Bulletin, 21, 125 - 135.

Tremblay, M. S., & Wilms, D. (2000). Secular trends in the body mass index of Canadian children. Canadian Medical Association Journal, 163, 1429 - 1433.

Vanden Auweele, Y., Bakker, F., Biddle, S., Durand, M., & Seiler, R. (1999) Psychology for physical educators. Champaign, IL: Human Kinetics.

Weiss, M. R. (1986). Self-esteem and achievement in children's sport and physical activity. In D. Gould, & M. Weiss (Eds.), Advances in pediatric sport sciences: Vol. 2, (pp. 87 - 119). Champaign, IL: Human Kinetics.

Weiss, M. R., & Bressan, E. S. (1985). Relating instructional theory to children's psychosocial development. Journal of Physical Education, Recreation and Dance, November/December, 34 - 36.

CHAPTER 3

Study 2: The Importance of Content-Relevance Evidence to Support the Construct Validity of the Children's Perceptions of Optimal Challenge Instrument

The use of self-report instruments or questionnaires is a commonly used method in psychological research. The assumption is that by responding to a number of pre-determined, closed-ended items which "capture" the phenomenon in question, that psychological insight will be gained. Once the items have been developed, a series of statistical procedures are often used to provide validation evidence that the instrument does indeed tap into the intended construct. Despite this seemingly carved out path for instrument construction, the importance of the initial development of the items is often overlooked. This concern is perhaps even more prevalent in psychological research conducted with children due to the tendency to simply change the wording (Brustad, 1998; Stone & Lemanek, 1990). Developmental psychology has come a long way in demonstrating that children are not miniature adults (Flannery, 1990). Therefore, when developing self-report instruments that are intended to capture a psychological construct with children, one needs to be diligent in the initial stages to ensure that the items provide content validation evidence which is consistent with children's developmental characteristics and the way children experience the construct.

Classical test score theory (Crocker & Algina, 1986) has proposed that an individual's observed score is the sum of his/her true score plus measurement error. Thus, the less measurement error, the closer the observed score is to the true score. Although one can not say that if something is valid it is true, validity does allow researchers to say that if the premise leads to the conclusion, then the premise is a valid argument (Mautner, 1997). In the case of self-report instruments, one way to control for potential measurement bias is to ensure that the items do indeed match the construct for which they were intended to measure.

Traditionally, the validity of an instrument intended to tap into a theoretical construct was determined by providing separate pieces of evidence pertaining to content, criterion-related, and construct validity. Recent thinking, however, has suggested that validity be considered from a unitary perspective rather than separate and distinct approaches (Angoff, 1988; Cronbach, 1988; Messick, 1989). The argument put forth for this rationale is that various sources of evidence are needed in order to provide support for inferences of scores produced by an instrument. However, in the development of the instrument, one must be sure that the items that have been developed have some evidence pertaining to their content validity in order to avoid later complications.

Content validation processes are concerned with "... showing how well the content of the test samples the class of situations or subject matter about which conclusions are to be drawn" (Messick, 1989, p. 16). Often, this content is obtained by going to primary and secondary sources. The primary sources are those who actually experience the construct in question. In the case of conducting psychological research with children, for example, investigators must go directly to children to understand how they experience and describe the construct. This will help in the writing of developmentally appropriate items. Secondary sources can also be utilised in the content validation process. Going to previous research and using items from existing instruments that have demonstrated adequate psychometric properties are other methods of obtaining content validity evidence. Another method, yet often overlooked and under-used, is to ask experts to judge the content relevance of the items (Dunn, Bouffard, & Rogers, 1999). Assuming that the judges are experts who are able to evaluate the relevance of the items relative to the construct for which they were intended, content validation evidence can be presented.

One construct that has been relatively under-studied in the area of physical activity psychology research with children is optimal challenge. This construct has been suggested to have an important influence on children's intrinsic motivation during physical activity participation (Csikszentmihalyi, 1975; Deci & Ryan, 1985; Harter, 1978). For example, when children perceive an activity to be too easy or too hard, it may lead to boredom or anxiety respectively. Alternatively, when children perceive their skills to be equal to the challenge of the activity, then they are likely to be optimally challenged. It has been hypothesised by many (e.g., Biddle & Chatzisarantis, 1999; Csikszentmihalyi, 1990; Reeve, 1996; Stipek, 1998; Weiss & Bressan, 1985) that when children are optimally challenged, they are intrinsically motivated to do the activity due to the positive subjective experience that is produced as a result of enhanced competence and enjoyment (Danner & Lonky, 1981).

Despite these hypotheses and the importance that perceptions of optimal challenge seem to have on children's intrinsic motivation, there is a lack of psychometric evidence to support using current self-report instruments with children. A previous study provided evidence that children describe, experience, and prefer physical activity experiences in relation to optimal challenge (Mandigo, 2001). The only instruments currently available are either unidimensional (e.g., Csikszentmihalyi & Csikszentmihalyi, 1988; Delignières, 1999; Harter, 1978), intended for adults (e.g., Csikszentmihalyi & Larson, 1987; Jackson & Marsh, 1996), or objective rather than subjective (e.g., Danner & Lonky, 1981; Sarrazin & Famose, 1999). For example, Harter (1978) used length of anagram (3-6 letters) and a 4 point self-report rating of the difficulty of the task (1 = very easy; 4 = very hard) as indicators of optimal challenge. She then inferred from high levels of affect that children were optimally challenged when they worked on 4 or 5 letter anagrams and rated the task somewhere between 1 and 4 on the difficulty scale. Nowhere in the study did children actually report whether they were optimally challenged. In addition, although Jackson and Marsh's (1996) Flow State Scale contains a "skill/ challenge balance" subscale, the items were developed for elite athletes, not children. Finally, concerns have been raised regarding the skill and challenge measures contained in one of the most common instruments to measure optimal challenge; the Experience Sampling Form (Csikszentmihalyi & Larson, 1987). Traditionally, individuals have used this instrument to rate how challenged they felt during an activity and how much skill they had to do the activity. Like optimal challenge, participants are considered to be in flow when they perceive their skill level and the challenge of the activity to be equal and above average. However, many researchers have expressed trepidation over the use of the one item, unidimensional measures of skill and challenge (Voelkl & Ellis, 1998). The purpose of this study, therefore, was to provide the next step towards creating a developmentally appropriate self-report instrument by presenting content validity evidence for items intended to capture children's perceptions of optimal challenge. The items have originated from a study of children's comments and perceptions of optimal challenge (Mandigo, 2001).

Method

Due to the absence of a psychometrically sound self-report instrument that provides information on children's perceptions of optimal challenge, items for a new self-report instrument had to be developed from a variety of sources. The primary source of item development was obtained from a content analysis of interviews with 27 children about their experiences in various types of physical activities (Mandigo, 2001). Children's descriptions, outcomes, and preferences for concepts embedded within the construct of optimal challenge (e.g., skill and challenge are balanced or unbalanced) were used. The items that were generated from the content analysis were used to reflect how children would describe and experience different levels of optimal challenge. The use of participant-generated items is an important first step to improving the relevance and hence content validity of the instrument (Dunn et al., 1999). Relevant research from the children's motivation literature was also reviewed to provide additional items or to clarify newly developed items.

Together, this information helped to facilitate the development of items for a preliminary draft of the Children's Perceptions of Optimal Challenge Instrument (CPOCI). Csikszentmihalyi's (1975) flow model was used as the theoretical underpinning along with children's interpretations of words associated with this model (e.g., challenge, skill, skill/ challenge balance) to develop a table of specifications (Rust & Golombok, 1989). According to this theory, individuals are optimally challenged when they perceive their skill to be equal to the challenge of the activity. Alternatively, individuals are not optimally challenged when these perceptions are not balanced. In order to capture the entire construct of optimal challenge based on Csikszentmihalyi's (1975) flow model, it is important to develop content areas that capture the wide range of potential perceived skill and challenge ratios. As a result, three separate content areas were developed:

i) Skill Greater than Challenge: participants perceive their skills to be greater than the challenge of the activity;

ii) Challenge Greater than Skill: participants perceive their skills to be less than the challenge of the activity; and,

iii) Skill Equals Challenge: participants perceive that their skills are equal to and balanced with the challenge of the activity.

The items (n = 27) were then generated by going back to the content analysis conducted in the previous study (Mandigo, 2001) and using terminology that was consistent to the way children would describe and experience these three content areas. The generated items were then examined by the lead researcher, members of his doctoral supervisory committee (n = 4), and a book editor from a children's educational publishing company to ensure the items met the recommended standards for a good item (DeVillis, 1991) and were at an appropriate readability level for children as young as seven years of age. For example, lengthy items were avoided and children's language from previous interviews were used as much as possible to avoid problems with reading level. Potential "double barrelled" items were also screened out to ensure an item did not endorse more than one idea and items which reflect "double negative" terms were reworded to avoid confusion. Following this stage, the items were reviewed by a panel of expert judges with respect to their relevance to the construct of optimal challenge for children.

Participants

A total of 15 potential experts who had either a teaching and/or research background in the area of intrinsic motivation for children were contacted regarding their interest in participating in the study. To be qualified as an expert judge, the participant had to be on faculty at a post-secondary institution, have a minimum of a Masters degree, and have an understanding of research pertaining to children's motivation in physical activity. Judges also represented various theoretical perspectives on motivation (e.g., Self-Determination Theory, Self-Regulation; Goal Orientation; Competence Motivation). A total of 11 experts agreed to participate in the study. After several follow-up contacts, 9 judges (3F; 6M) completed and returned the rating forms.

Instruments

Experts were asked to rate the relevance of the items using the Item Content Relevance Form (ICRF). The ICRF requires each judge to evaluate the relevance of each item to each content area (Dunn et al., 1999). Judges used a 5-point Likert-type scale (1 = poor match; 5 = excellent match) to rate the match of an item to each of the three content areas. The judges then used the ICRF to evaluate how well one item matches the content area of skill equals challenge, skill greater than challenge, and challenge greater than skill. An item would be deemed relevant should judges provide a high-match rating when comparing it to the target content area (i.e., item is from the content area) and a low-match rating when comparing it to the non-target content area (i.e., item is not from the content area). For example, judges rated how well the first item (i.e., it was in the middle of easy and hard) reflected each of the content areas (i.e., skill equals challenge, challenge equals skill, and skill equals challenge). Because this item was written to be included in the challenge equals skill subscale, it would be hypothesised that judges would rate this item high for this content area and low for the other two (i.e., skill and challenge not equal) content areas. This approach reduced the amount of rating bias due to judges being "blind" to the intended item-content matches (Dunn et al., 1999). Judges were also encouraged to provide written comments.

Procedures

Potential judges were solicited by phone, personal contact, and/or electronic mail. Those who agreed to participate in the study were asked to indicate whether they wished to receive the ICRF by fax, mail, or as an electronic mail attachment. Each judge was sent an information letter and informed consent form that they were asked to read, sign, and return. Judges had the option of sending the ICRF back with the consent form, or to send it back by fax, mail, or as an electronic mail attachment.

After receiving the judges' responses, mean item content-relevance ratings, Aiken's (1985) content-validity coefficient (*V*), and Cohen's (1977) effect size (ES) index for dependent means were calculated to determine the relevance of the items across the three content areas. Aiken's (1985) *V* coefficient provides information on the significance of judges' ratings for the content area that each item was designed to measure. It is determined by $V = S/[n(c-1)]$ where:

S = Σs;

s = r - lo;

r = judge's validity rating;

lo = lowest possible rating;

n = number of judges; and

c = number of successive integers on the rating scale (i.e., a 5-point scale).

The *V* coefficient can range from 0 to 1. A value of 1.0 indicates that all judges gave an item the highest possible score on the rating scale while a value of 0 indicates that all judges gave the item the lowest possible rating score. The statistical significance of *V* was then determined by consulting Aiken's (1985) right-tailed binomial probability table.

In order to determine whether other non-keyed items were viewed as being relevant for the content areas they were not intended to measure, Cohen's (1977) ES index for dependent means (d_z') was computed using the following formula:

$$d_z' = \frac{\bar{X} - \bar{Y}}{\sqrt{\sigma^2_x + \sigma^2_y - 2\, r_{xy}\, \sigma_x \sigma_y}}$$

d_z' = effect size index for dependent means

$\bar{X}$ = mean rating for variable X

$\bar{Y}$ = mean rating for variable Y

σ^2_x = variance of variable X

σ^2_y = variance of variable Y

$2\, r_{xy}\, \sigma_x \sigma_y$ = covariance of X.Y.

This measure is analogous to planned contrasts whereby the mean content-relevance score for the keyed item is compared to the mean content-relevance score for the non-keyed item (Dunn et al., 1999). Using Cohen's (1977) guidelines, a d_z' of .80 or greater is considered to show a large ES while a d_z' of .50 - .79 is considered to represent a moderate ES.

Results

Table 3.1 presents the mean keyed response (i.e., the item is part of the intended content area) and corresponding standard deviation for each item. Aiken's *V* coefficient is also presented and as indicated, 21 of the 27 items were significant at the .05 level. The mean of all content relevance values across all 27 items was equal to 4.07 (SD = .49). When all of the items that were not significant were taken out, the mean content relevance rating of the remaining keyed items was equal to 4.28 (SD = .26). This represents a "Very Good Match" when using the Likert-scale anchors from the ICRF.

An indication of the dimensionality of the items is presented in Table 3.2. Using Cohen's (1977) guidelines for establishing effect size, only two of the fifty-four planned contrasts were below the .80 level for a large effect size. This suggests that the majority of keyed items were significantly different from the other two content areas that they were not intended to measure. As such, it can be argued that the items within each content area do not provide an indication of a participant's perception of skill equals challenge for the other two content areas.

Judges were also asked to provide comments regarding their thoughts on the items. The majority of comments came from keyed items that were rated low by the judges. For example, judges raised concern about item four [I did not know how to do it] which was to be included in the Challenge > Skill content area. Two judges in particular were concerned that this item was tapping into a different construct all together; that being knowledge and understanding. As such, this item received a low rating and a non-significant *V* coefficient. Item 10 [It made me feel good to do something challenging] was perceived as reflecting an outcome of skill equals challenge (i.e., positive affect) rather than the actual state of being optimally challenged. As such, this item had the lowest mean rating and *V* coefficient of all the keyed items. Similar concerns were raised over item 13 [It was a fun challenge that I wanted to keep doing] which was perceived as compounding challenge and fun (an outcome) together.

Concern was also raised by judges about the wording of Item 5 [I had enough skill to do a good challenge]. Although children in a previous study distinguished between a good and bad challenge (Mandigo, 2001), perhaps the wording of the question was not as clear as it could have been. For example, three judges expressed concern over how a child can "do" a good challenge. Therefore, this item was eliminated and item 14 [It was a good challenge that I could do] was retained instead due to its high relevance rating ($\underline{M}$ = 4.00; $\underline{SD}$ = .87) and significant *V* coefficient of .75 ($\underline{p}$< .05).

Although items 3 [I could not do it because it was too hard] and 19 [It was really simple for me to do] received a high mean rating by the judges, both received a "Poor Match" by two and one judge respectively. This represents the lowest rating an item can receive. Therefore, caution may need to be taken if these items are to be included for future uses.

Discussion

The results from the current study are valuable in developing validation evidence for a self-report instrument intended for children. Asking experts in the field to rate the relevance of items provides an indication of their content validity. The results are even more meaningful considering that the majority of the items were developed with the words children used to describe their perceptions of optimal challenge (Mandigo, 2001) and supported by the expert judges. As the results indicate, the judges agreed with the majority of the items (i.e., 21/27) that were generated from the previous research. Hence, content validity evidence pertaining to the relevance of the items has been provided from both primary and secondary sources. Although evaluating the validity of any psychological measure is an on-going process (Flannery, 1990), researchers can now proceed with confidence that the content validity of the three subscales has been established through various sources due to their match to the three content areas.

Future research will need to examine both the internal and external structures of the instrument in order to provide further validation evidence. For example, idiographic validation procedures can be used to ensure that the interpretations derived from the items reflect the construct of optimal challenge at an individual level. Further validation procedures can also use nomothetic methodologies that determine the validity of the results at the group level. If, after using both approaches, the interpretations from the instrument still support the validity of the instrument, then researchers can continue refining the instrument with confidence that the items contain minimal measurement error and are in fact relevant to the three subscales developed to measure different levels of optimal challenge. Although only one subscale is representative of being optimally challenged (i.e., skill equals challenge), the other two subscales provide information on the degree to which individuals are not optimally challenged. Future research should assess the overall representativeness (Crocker & Algina, 1986) of the three subscales and their relationship to the overall construct of optimal challenge (i.e., skill equals challenge, skill greater than challenge, challenge greater than skill). This was not done in this study and will be extremely valuable further on in the validation process and during intervention studies.

Once a means to measure the overall construct of optimal challenge has been established, researchers will be in a better position to conduct experimental studies which examine the influence of situational (e.g., use of rewards, teaching/ coaching style, use of feedback) and dispositional (e.g., goal orientation, perceived competence, gender) factors on children's perceptions of optimal challenge. For example, Harter (1978) has found the use of rewards tend to undermine children's preference for harder tasks. Sarrazin and Famose (1999) have also suggested that a child's goal orientation has an influence on the amount of motivation across different levels of optimal challenge. However, due to the type of environment and the way in which optimal challenge was operationalised in previous research, the results may not be generalizable to children's perceptions of optimal challenge in physical activity environments. If a psychometrically sound instrument is available, then further insight can be gained on how to structure a physical activity environment for children that will motivate them to be active. The results from this study have laid the ground work in developing such an instrument and are a step in the right direction down the validation road towards creating a developmentally appropriate and psychometrically sound self-report instrument of children's perceptions of optimal challenge.

Table 3.1.

Mean Item-Content Relevance Ratings and *V* Coefficient for each Item Based on the Content Domain it was Originally Designed to Measure

	Item Description	Content Domain	Mean	SD	*V*
1.	It was in the middle of easy and hard	S=C	4.00	.71	.75*
2.	I could do it really easily	S > C	4.44	.53	.86**
3.	I could not do it because it was too hard	C > S	4.00	1.73	.75*
4.	I did not know how to do it	C > S	3.44	1.13	.61
5.	I had enough skill to do a good challenge	S=C	3.25	1.39	.56
6.	I had more skill than I needed to do it	S > C	4.22	.97	.81**
7.	I needed more practice to be able to do it well	C > S	3.22	1.09	.56
8.	I needed more skill to do the activity	C > S	4.33	1.12	.83**
9.	I was not very good at it	C > S	4.00	.87	.75*
10.	It made me feel good to do something challenging	S=C	2.77	.97	.44
11.	It should be made a lot harder	S > C	3.56	1.33	.64
12.	I wanted to quit because it was too hard	C > S	4.78	.44	.94**
13.	It was a fun challenge that I wanted to keep doing	S=C	3.67	1.22	.67
14.	It was a good challenge that I could do	S=C	4.00	.87	.75*
15.	It was easy because I had done it lots of times before	S > C	4.22	.97	.81**
16.	It was challenging, but I could do it well	S=C	4.11	1.05	.78*
17.	It was not too easy and not too hard	S=C	4.11	.78	.78*
18.	It was not very challenging for me to do	S > C	4.00	1.00	.75*
19.	It was really simple for me to do	S > C	4.11	1.27	.78*
20.	It was so easy, I did not want to do it any more	S > C	4.44	.73	.86**
21.	It was too challenging for my skill level	C > S	4.55	.73	.89**
22.	It was too hard for me to do	C > S	4.44	.53	.86**
23.	My skills were a lot higher than the challenge	S > C	4.44	.73	.86**
24.	My skills were equal to the challenge of the activity	S=C	4.44	.53	.86**
25.	The challenge was higher than what I could do	C > S	4.00	1.00	.75*
26.	The challenge was perfect for my skill level	S=C	4.67	.71	.92**
27.	It was so easy it was boring	S > C	4.67	.50	.92**

*$p < .05$; ** $p < .01$

NOTE: S=C: Skill equals Challenge; C > S: Challenge is higher than skill; S > C: Skill is higher than challenge.

Table 3.2.

Mean Content-Relevance Scores and Mean-Difference Effect Sizes for Ratings

Item	Mean Content Ratings			Effect Sizes for Planned Mean Contrasts	
	Skill = Challenge (i)	Skill > Challenge (ii)	Challenge > Skill (iii)	Contrast 1	Contrast 2
1. [S=C]	**4.00**	1.11	1.22	[i-ii] 3.70	[i-iii] 2.86
2. [S > C]	1.33	**4.44**	1.00	[ii-i] 3.35	[ii-iii] 6.54
3. [C > S]	1.00	1.67	**4.00**	[iii-i] 1.73	[iii-ii] 0.84
4. [C > S]	1.00	1.11	**3.44**	[iii-i] 2.16	[iii-ii] 1.76
5. [S=C]	**3.25**	1.63	1.63	[i-ii] 1.15	[i-iii] .84
6. [S > C]	1.22	**4.22**	1.00	[ii-i] 2.68	[ii-iii] 3.31
7. [C > S]	1.89	1.00	**3.22**	[iii-i] .77	[iii-ii] 2.03
8. [C > S]	1.33	1.00	**4.33**	[iii-i] 1.90	[iii-ii] 2.98
9. [C > S]	1.11	1.00	**4.00**	[iii-i] 3.11	[iii-ii] 3.46
10. [S=C]	**2.78**	1.44	2.00	[i-ii] 1.09	[i-iii] 0.65
11. [S > C]	1.00	**3.56**	1.00	[ii-i] 1.92	[ii-iii] 1.92
12. [C > S]	1.11	1.00	**4.78**	[iii-i] 5.19	[iii-ii] 8.57
13. [S=C]	**3.67**	1.44	1.22	[i-ii] 1.42	[i-iii] 1.62
14. [S=C]	**4.00**	1.56	1.00	[i-ii] 1.62	[i-iii] 3.46
15. [S > C]	1.33	**4.22**	1.00	[ii-i] 2.74	[ii-iii] 3.31
16. [S=C]	**4.11**	2.00	1.44	[i-ii] 1.07	[i-iii] 1.69
17. [S=C]	**4.11**	1.00	1.00	[i-ii] 3.98	[i-iii] 2.28
18. [S > C]	1.11	**4.00**	1.00	[ii-i] 2.28	[ii-iii] 3.00
19. [S > C]	1.11	**4.11**	1.33	[ii-i] 1.90	[ii-iii] 1.25
20. [S > C]	1.00	**4.44**	1.00	[ii-i] 4.75	[ii-iii] 4.74
21. [C > S]	1.00	1.00	**4.56**	[iii-i] 4.89	[iii-ii] 4.89
22. [C > S]	1.00	1.00	**4.44**	[iii-ı] 6.54	[iii-ii] 6.54
23. [S > C]	1.00	**4.44**	1.00	[ii-i] 4.74	[iii-ii] 4.74
24. [S=C]	**4.44**	1.44	1.00	[i-ii] 3.00	[i-iii] 6.54
25. [C > S]	1.11	1.00	**4.00**	[iii-i] 2.28	[iii-ii] 3.00
26. [S=C]	**4.67**	1.44	1.00	[i-ii] 2.48	[i-iii] 5.19
27. [S > C]	1.00	**4.67**	1.00	[ii-i] 7.33	[ii-iii] 7.33

Note: Means for keyed items are bolded.

References

Aiken, L. R. (1985). Three coefficients for analysing the reliability and validity of ratings. Educational and Psychological Measurement, 45, 131 - 142.

Angoff, W. H. (1988). Validity: An evolving concept. In H. Wainer & H. I. Braun (Eds.), Test validity (pp. 19 - 32). Hillsdale, NJ: Laurence Erlbaum.

Biddle, S. J. H., & Chatzisarantis, N. (1999). Motivation for a physically active lifestyle through physical education. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand, & R. Seiler (Eds.), Psychology for physical educators (pp. 5 - 26). Champaign, IL: Human Kinetics.

Brustad, R. J. (1998). Developmental considerations in sport and exercise psychology measurement. In J. Duda (Ed.), Advances in sport and exercise psychology measurement (pp. 461 - 470). Morgantown, WV: Fitness Information Technology.

Cohen, J. (1977). Statistical power analysis for the behavioural sciences (2nd ed.). New York: Academic.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Philadelphia, PA: Harcourt Brace Jovanovich.

Cronbach, L. J. (1988). Five perspectives on validity argument. In H. Wainer, & H. I. Braun (Eds.), Test validity (pp. 3 - 17). Hillsdale, NJ: Laurence Erlbaum.

Csikszentmihalyi, M. (1975). Beyond boredom and anxiety. San Francisco, CA: Jossey-Bass.

Csikszentmihalyi, M. (1990). Flow. New York: Harper & Row.

Csikszentmihalyi, M., & Csikszentmihalyi, I. (1988). Introduction to Part IV. In M. Csikszentmihalyi, & I. Csikszentmihalyi (Eds.), Optimal experience. Psychological studies of flow in consciousness (pp. 251 - 265). Cambridge, UK: Cambridge University Press.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. The Journal of Nervous and Mental Disease, 175, 526 - 536.

Danner, F. W., & Lonky, E. (1981). A cognitive-developmental approach to the effects of rewards on intrinsic motivation. Child Development, 52, 1043 - 1052.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behaviour. New York: Plenum.

Delignières, D. (1999). Perceived difficulty, resources investment, and motor performance. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand, & R. Seiler (Eds.), Psychology for physical educators (pp. 211 - 232). Champaign, IL: Human Kinetics.

DeVillis, R. F. (1991). Scale development: Theory and application. Newbury Park, CA: Sage.

Dunn, J. G. H., Bouffard, M., & Rogers, W. T. (1999). Assessing item content-relevance in sport psychology scale construction research: Issues and recommendations. Measurement in Physical Education and Exercise Science, 3, 15 - 36.

Flanery, R. C. (1990). Methodological and psychometric considerations in child reports. In A. M. La Greca (Ed.), Through the eyes of the child (pp. 57 - 108). Needham Heights, MA: Allyn & Bacon.

Harter, S. (1978). Effectance motivation reconsidered. Human Development, 21, 34 - 64.

Jackson, S. A., & Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The flow state scale. Journal of Sport and Exercise Psychology, 18, 17 - 35.

Mandigo, J. (2001). Listening to the voices of children: Taking the first steps in the development of a self-report instrument of children's perceptions of optimal challenge. Unpublished manuscript, University of Alberta, Edmonton, Alberta, Canada.

Mautner, T. (1997). Dictionary of philosophy. London, UK: Penguin.

Messick, S. (1989). Validity. In R .L. Linn (Ed.), Educational measurement (3rd ed.) (pp. 13 - 103). New York: American Council on Education.

Reeve, J. (1996). Motivating others: Nurturing inner motivational resources. Needham Heights, MA: Allyn & Bacon.

Rust, J., & Golombok, S. (1989). Modern psychometrics: The science of psychological assessment. London, UK: Routledge.

Sarrazin, P., & Famose, J. (1999). Children's goals and motivation in physical education. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand, & R. Seiler (Eds.), Psychology for physical educators (pp. 27 - 50). Champaign, IL: Human Kinetics.

Stipek, D. (1998). Motivation to learn. From theory to practice. Boston, MA: Allyn & Bacon.

Stone, W. L., & Lemanek, K. L. (1990). Developmental issues in children's self-reports. In A. M. La Greca (Ed.), Through the eyes of the child (pp. 18 - 56). Needham Heights, MA: Allyn & Bacon.

Voelkl, J. E., & Ellis, G. D. (1998). Measuring flow experiences in daily life: An examination of the items used to measure challenge and skill. Journal of Leisure Research, 30, 380 - 389.

Weiss, M. R., & Bressan, E. S. (1985). Relating instructional theory to children's psychosocial development. Journal of Physical Education, Recreation and Dance, November/December, 34 - 36.

CHAPTER 4

Study 3: Using Children's Similarity Ratings as a Source of Construct Validity Evidence

Developing self-report instruments for children that produce accurate interpretations of a psychological construct is not an easy task. Children are not miniature adults and hence, the assumptions underlying psychological theories and measurement instruments intended for adults are not always transferable. Developmental research has illustrated that children's interpretations and understanding of the world is often vastly different to that of adults (Crain, 1992; Miller, 1993). Differences due to physical, cognitive, and affective characteristics of children have a profound influence on research pertaining to children's psychological experiences. In order to ensure that the observed score provided by the scores from psychological instruments have minimal amounts of measurement bias, validation evidence must first be provided such that the items which make up the instrument accurately reflect how children experience and interpret the construct (Brustad, 1998).

In psychometric research, evidence of an instrument's validity can be provided through various techniques. Validity has been defined as: "... an integrated evaluative judgement of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment" (Messick, 1989, p. 13). Although one can not say that if something is valid it is true, validity does allow researchers to say that if P, the premise, leads to C, the conclusion, then P is a valid argument (Mautner, 1997). In the case of self-reports, various sources of evidence are collected in order for researchers to present the case that the instrument is a reflection of the construct. When conducting psychological research with children, questions surrounding the validity of an instrument must not only include traditional sources of validation evidence, but questions concerning the children's developmental characteristics and capabilities must also be addressed. Children must be capable of understanding the items on the self-report instrument and the construct must also be defined in a manner that is consistent to the way it would be experienced and described by children. It can not be assumed that just because a construct has been suggested to exist in the "adult world" that it is experienced or described in exactly the same way for children. Thus, when developing self-report instruments for children, it is extremely important to go directly to the children in order to ensure that the items that are meant to represent the construct in question are in fact, interpreted that way by the children. Failure to do so may result in inappropriate amounts of measurement error such as construct-irrelevance variance and/or construct-under-representation variance which can cause individuals' observed score to be further away from their true score (Crocker & Algina, 1986).

Often, however, researchers follow a common pattern of instrument development whereby the items are created based on previous research or experience, and then administered to a large group of participants without previously consulting with them. The items are then factor analysed and should the items form a coherent factor structure in the manner in which they were intended and have a high level of internal consistency, it is argued that the items are a valid and reliable representation of the construct for which they were intended. Although this may be the most efficient method to instrument development, it does not always ensure that the items have good content validity as defined as: "... evidence that the test items, tasks, or procedures are representative of a previously defined universe or domain of content" (Flannery, 1990, p. 69). To do this, it is important to go directly to the primary source for whom the instrument is intended.

Validation is a never ending process (Messick, 1989) and thus, various sources of evidence must be provided to ensure the scores provide an accurate representation of the intended construct. This study is the third step in a series of studies aimed at the development of items that will accurately determine whether or not children have been optimally challenged during participation in physical activity.

The construct of optimal challenge was chosen for several reasons. First, it has been recognised in the motivation literature as playing an important role in motivating individuals. Organismic theories of motivation argue that when children perceived their skills to be balanced with their perceptions of the challenge of the activity they are likely to be intrinsically motivated to participate due to the positive subjective states associated with this construct (Csikszentmihalyi, 1975; Deci, 1975; Harter, 1978). Alternatively when perceived skill and challenge are unbalanced, negative outcomes such as boredom and anxiety often arise. This type of approach is embedded within a comprehensive model of optimal challenge. Secondly, the importance of understanding children's optimal challenge has been identified within the participation and attrition motivation literature (e.g., Weiss & Bressan, 1985). Several studies have shown that when children are not optimally challenged (e.g., are bored or frustrated) this can often lead to withdrawal. When children are being challenged and having fun (i.e., perceptions of skill are equal to perceptions of challenge [Petlichkoff, 1992]), they are more likely to continue participation (e.g., Gould, Feltz, & Weiss, 1985; McCullagh, Matzkanin, Shaw, & Maldonado, 1993). Finally, despite these arguments and the importance that this construct has on motivating children to be physically active (e.g., Biddle & Chatzisarantis, 1999; Rowley, 1996; Weiss & Bressan, 1985), there does not exist a psychometrically sound and developmentally appropriate instrument for measuring optimal challenge of children. Therefore, if a psychometrically sound self-report instrument can be developed that will determine the degree to which children are optimally challenged during a physical activity or not, then researchers may better understand the situational factors and dispositional variables that influence children's motivation.

The purpose of this study was to determine the relevance of items by having children rate the similarities and differences between items intended to measure children's perceptions of optimal challenge. The items used in this study were initially developed following two previous studies. In study one (Mandigo, 2001a), children were interviewed about their experiences in physical education and physical activity as it related to various levels of skill and challenge. Based on these interviews, items that reflected different levels of optimal challenge (i.e., skill equals challenge, challenge greater than skill and skill greater than challenge) were developed. The relevance of these items was investigated in study two (Mandigo, 2001b). Expert judges in the field were asked to rate the relevance of the generated items as they pertained to different levels of optimal challenge. Items representing three subscales were created to provide an overall measure of the degree to which an individual is optimally challenged. The judges rated 21 of the 27 items in the initial pool to be relevant. Upon further inspection of the remaining

21 items, two items had a least one judge rate them as being a "Poor Match" and were thus excluded. A third item [It was not very challenging for me to do] was deleted because it had the lowest mean rating left of the remaining seven items for the "Skill Greater than Challenge" content area. This was done so that equal numbers (n= 6) of items per content area were included. The remaining 18 items represented the best six matches for each of the content areas based on the mean and standard deviations of the judges' ratings. Although this may seem like a trivial reason, one of the goals in developing the instrument was to make it concise so that children, who often have short attention spans, can answer the items in a brief period of time. Therefore, for the sake of parsimony and practicality, these 18 items were included. The current study sought to go back to the children to ensure that the remaining items were relevant to how children would describe different degrees of optimal challenge as a general construct.

Method

Participants

A total of 15 (4M; 11F) participants took part in this study. Grades 2 (n = 7) and 3 (n = 8) children were recruited from two different elementary public schools. Ten of the participants were from a school located in a western Canadian university setting and five participants were from a "typical" elementary school located in an urban, western Canadian city. All participants and their parents provided informed consent prior to the start of the data collection.

Procedure

Upon entering the interview room, students were told the rights that they had as a participant in the study. They were told that this was not a test and that their performance was not being marked and that only the researcher would see their data. They were also told that they could stop participating at any time without consequence and were encouraged to ask any questions. They were then asked if they were ready to participate in a couple of practice tasks before doing the experimental task for the study.

Practice tasks. Practice tasks that are fairly easy and introduce participants to the rating scale are an important first step when participants are unfamiliar with judging similarities between items (Schiffman, Reynolds, & Young, 1981). The first practice task required participants to judge the similarity between different colours. A coloured piece of construction paper was placed in the middle of a game board (see Figure 4.1). This game board consisted of two sides: Similar and Not Similar. The Similar side was divided into Really Similar and Sort of Similar halves while the Not Similar side was divided into Really Not Similar and Sort of Not Similar. Participants were then given eight coloured pieces of construction paper and told to decide whether each colour in their hand was Similar or Not Similar to the colour in the middle of the game board. Once they made that decision, they were then asked to decide whether the colour was Really or Sort of Similar or Not Similar. Participants did this a few times until they felt comfortable doing the task.

The second practice task required participants to use the same procedure as above but with cards containing statements about different types of weather. These items were:

i) Today is a very sunny day;
ii) It is very bright outside today;
iii) It is raining a lot outside today;
iv) Today is a very wet day;
v) It is really windy outside today; and,
vi) The wind is very strong today.

Using the game board, participants were told to compare the meaning of the entire sentence to the meaning of one of the cards that had been placed in the middle. For example, the card "today is a very sunny day" was placed in the middle and all other statement cards were compared to it. After participants had compared each item to one another, they were asked to take the cards and sort them into three separate piles. Each pile was to be different from each other, but the cards in each pile were to mean the same thing. Every participant was able to group the cards into sunny, raining, and windy piles.

Experimental task. At the beginning of the experimental task, participants were given all 18 items (see Figure 4.2) which were typed on a laminated card, and asked to read through each one. A modified version of the Rotating Standard Method (Davison, 1983) was then used to obtain similarity and dissimilarity ratings between items. Although this method does not control for time or space effects, it does speed up the process and hence, control for children's short attention spans. A reference item was randomly chosen and placed in the middle of the game board. Participants were told that they would be comparing the meaning of all remaining cards to the reference item in the middle. It was pointed out to the participants that they would see similar words on the cards, but it was the meaning of the entire sentence in which they would be making their comparisons. Participants then decided whether the meaning of the card they had in their hand was Similar or Not Similar to the reference item. Once they made that decision, they would then decide whether the meaning was "Really Similar" or "Sort of Similar" if they decided the meaning was Similar or "Really Not Similar" or "Sort of Not Similar" if they decided the meaning was Not Similar. Paired comparisons were assigned a value from one to four where a one represented "Really Similar" and a four represented "Really Not Similar". Each card contained the item number on the back that only the researcher could see so he could record the score as the participants placed them in the squares. Once all of the remaining 17 items were compared to the first reference item, it was taken out and a new reference item was randomly chosen from the pile. The remaining items were compared to all but the first reference item to avoid redundancy. This resulted in a total of 153 paired comparisons for each participant.

The procedure indicated above was a modified version of the Rotating Standard Method (Davison, 1983) or what some call the Category Rating Technique (Schiffman et al., 1981). Traditionally, participants use paper and pencil techniques to indicate how similar or dissimilar they think two stimuli are by rating along a bipolar rating scale bounded by "not at all similar" to "very similar" descriptions. A value is then assigned based on where the participant rates the degree of similarity. The procedure used in this study was modified in order to make it as developmentally appropriate as possible and hence, control for potential task irrelevant sources of measurement error associated with procedures that are too difficult for children (Crocker & Algina, 1986). First, the game board "questionnaire" was used instead of a paper and pencil questionnaire to stimulate the children's interest and control for their short attention spans (Miller, 1993). Modifying the presentation of the task to one that seemed more like a puzzle or a game increased interest in the task and as a result, maintained the children's attention span for a longer period of time. Secondly, traditional "scaling" techniques use only the two most extreme descriptive anchors at either end of the bipolar scales. Participants are then asked to indicate their rating judgement somewhere along this line. For children, however, the lack of descriptive anchors along the scale may be too abstract. Based upon developmental literature, "... [children] require a concrete (tangible or visible) representation of relationships in order to effectively solve problems" (Brustad, 1998, p. 462). Although some researchers have expressed concerns over using more than two descriptive anchors when using the category rating technique (Schiffman et al., 1981), it was felt descriptive anchors were needed to provide concrete ratings when using this procedure with children. Based upon the strong psychometric properties found for Harter's (1985) Perceived Competence Scale for Children, the descriptive anchors of "Really" and "Sort of" that were used in that instrument were chosen for this study to allow children with concrete comparisons.

Once the participants had finished rating the pairs, they were then asked to take all of the cards and sort them into three piles in order to determine whether children agreed with the pre-determined content areas. Children were instructed that each pile was to be different from each other and the cards that were placed in each pile were to have the same meaning. Once all of the cards were sorted into piles, participants were asked to give each pile a name that would represent the meaning of the cards in the pile. The technique, known as the sorting technique (Kruskal & Wish, 1978) is analogous to a simplified factor analysis where items from similar factors load on each other. Although participants were told to group the cards into three piles, they were also told that they could make more piles if they felt it was necessary.

The procedures described above were pilot tested prior to use in the study. One female participant was taken through an initial protocol and asked to provide feedback on the task. She felt that the procedure was suitable and that the items could be read and understood by children as young as Grade 2.

Results

Using the ALSCAL procedure in SPSS, multidimensional scaling (MDS) analyses using Euclidean distance scaling models were conducted in provide a spatial representation of the data sets at the group and individual levels. By asking participants to directly assess the similarity between items, a dissimilarity matrix between items can be generated. This matrix is then analysed and a geometric configuration can be provided which clusters items rated to be similar together and separates items perceived to be different from each other. By using this information, a neighbourhood interpretation can be used to examine clusters that may contain items that have shared meanings (Kruskal & Wish, 1978). MDS also provides an indication of the dimensionality of the data without requiring large sample sizes. For each level of dimensionality, two goodness-of-fit indices are produced: stress and RSQ. Stress is defined as "... the square root of a normalised residual sum of squares" (Kruskal & Wish, 1978, p. 49). Stress values can range from 0 - 1.00 where a 0 indicates a perfect fit. Stalans (1997) recommended that small stress values range between 0 - .15. RSQ is defined as "... the proportion of variance of the transformed data that is accounted for by the resulting MDS solution" (Dunn, 1999, pp. 263-264). Thus, the higher the RSQ value, the more variance that is accounted for by the solution. RSQ values can also range from 0 - 1.00 where an RSQ value of 1.00 indicates a solution in which all variance is accounted for and thus provides a better fit. RSQ values between .80 and 1.00 are usually considered high (Stalans, 1997). However, caution must be taken when choosing the correct level of dimensionality. By its very nature, higher levels of dimensionality will produce lower stress values and higher RSQ ratings (Kruskal & Wish, 1978). The dimensionality that is ultimately chosen must be interpretable. A dimension that is not interpretable is likely not useful (Schiffman et al., 1981). Therefore, goodness-of-fit indices must be examined in conjunction with the interpretation of the data along some sort of theoretical perspective when choosing the appropriate level of dimensionality. For the purposes of this study, both the dimensionality of the solutions and the clustering of the items (i.e., neighbourhood interpretation) were used to understand as much of the data and results as possible (Kruskal & Wish, 1978). Due to the influence of potential dispositional variables that may affect children's preference for optimal challenge, each individual's dissimilarity matrix was analysed in addition to the overall group matrix in order to ensure that important and relevant information was not lost at the individual level (Bouffard, 1993).

Group Results: Category Rating Technique

Table 4.1 provides a summary of the proximities matrix that was created by taking the mean paired comparison ratings across thirteen of the fifteen individuals. Two participants chose to withdraw from this part of the study mid-way through their comparisons (likely because they were bored) and thus, their data was excluded. Although the goodness-of-fit values for the three dimensional solution (Stress = .10; RSQ = .95) are better, the two dimensional solution (Stress = .15; RSQ = .92) still provides adequate goodness-of-fit indices and provides the best theoretical interpretation. The one dimensional solution did not provide adequate goodness-of-fit indices (Stress = .30; RSQ = .85). Based upon several organismic theories of intrinsic motivation as a model (Csikszentmihalyi, 1975; Deci, 1975; Harter, 1978), Dimension 1 along the X-axis was labelled "Perceived Challenge/ Skill Balance" and Dimension 2 along the Y-axis was labelled "Enjoyment" (see Figure 4.3).

Items from the "Challenge Greater than Skill" subscale clustered high on the perceived challenge/ skill balance dimension (i.e., suggesting an imbalance where challenge is higher than skill) and low on the enjoyment dimension. Items from the "Skill Greater than Challenge" subscale clustered low on the perceived challenge/ skill balance dimension (i.e., suggesting an imbalance, but in the opposite direction to the challenge grater than skill items with skill higher than challenge) and low on the enjoyment dimension. Alternatively, items from the "Skill Equals Challenge" subscale tended to cluster in the middle (i.e., challenge and skill are equal) of the perceived challenge/ skill dimension and high on the enjoyment dimension. These results support the theoretical interpretation of children's optimal challenge based upon an organismic perspective of motivation (Deci, 1975; White, 1959). It appears that when children perceived their skills to be balance with the challenge of an activity (i.e., challenge = skill), it produced a high level of enjoyment. This hypothesis has been supported by previous research with children (Harter, 1974; 1978; Mandigo & Couture, 1996).

The data were also examined to look for potential differences in interpretations between the Grade 2 and Grade 3 participants. At the two dimensional solution, the goodness-of-fit indices were adequate for both the Grade 2 (Stress = .22; RSQ = .80) and

Grade 3 (Stress = .15; RSQ = .91) participants. As Table 4.2 indicates, the stimulus co-ordinates for the two dimensional solution showed similar geometric configurations for both grades when compared to the entire group.

Individual Results: Category Rating Technique

Individual proximity matrices based on categorical ratings were examined using ALSCAL MDS for the 13 individuals who completed the rating portion of the study. Table 4.3 provides a summary of individual goodness-of-fit values for each of these individuals. Although the stress and RSQ values were better at the three dimensional solution, the best theoretical interpretation of the results was at the two dimensional solution. For the majority of the participants, the two dimensional solution provided adequate goodness-of-fit values and the dimensions could be interpreted the same as the interpretation provided at the group level. However, participant 10 had poor goodness-of-fit values. This may be attributed to this participant's inability to focus for a long duration of time during the task. After examining this individual's geometric configuration, this data was taken out of the group solution due to the lack of a consistent pattern among the items. Although the goodness-of-fit values for the two dimensional solution at the group solution remained adequate without data from participant ten (Stress = .14; RSQ = .93), researchers may still wish to use discretion when using this method with children who have a difficult time remaining focused for extended periods of time.

Despite the individual variations due to potential dispositional differences, a number of individual profiles fit the hypothesis generated from the group solution. For example, Figure 4.4 provides an example of how all but item 14 (It was so easy it was boring) fit this hypothesis for participant number six. Across most participants, items from the same subscale tended to cluster together and to cluster along the hypothesised dimension at the predicted location. However, the amount of variation and stimulus configuration differed slightly among participants. Therefore, it is important to try and find out "how" participants were interpreting the meaning of the items by using a method called the "sorting technique".

Group Results: Sorting Technique

Individuals used the sorting technique to group similar items into the same piles. Schiffman et al. (1981) suggested that by assigning a "zero" to items grouped in the same pile and a "one" to items grouped in different piles, individual matrices then can be summed together to produce an overall group proximity matrix. This matrix was then analysed using ALSCAL MDS. Figure 4.5 provides the results of the two dimensional solution produced by this procedure. Again, although the goodness-of-fit values were slightly better at the three dimensional solution (Stress = .07; RSQ = .98), the two dimensional solution (Stress = .09; RSQ = .97) fit the best according to theoretical interpretations and with participants' descriptions of the groupings. The one dimensional solution did not provide adequate goodness-of-fit values (Stress = .09; RSQ = .97). The stimulus co-ordinates for the two dimension solution are presented in Table 4.4.

Similar to the categorical rating technique, items from the Challenge Greater than Skill subscale clustered high on the perceived challenge/skill dimension and low on the enjoyment dimensions while the Skill Greater than Challenge items clustered low on the perceived challenge/ skill dimension and low on the enjoyment dimension. As was expected, the Challenge Equals Skill items clustered in the middle of the skill/ challenge balance dimension and high on the enjoyment dimension. These interpretations are confirmed by the labels participants assigned to the groups. Most participants assigned the labcls of "Too Hard", "Too Easy" and, "In the Middle of Easy and Hard" to the groupings. After approximately seven participants had completed the study, it started to become apparent that a two dimensicnal solution may be the best theoretical representation of the data. Therefore, once participants had finished sorting the items into piles and naming the piles, the remaining eight were asked which one they enjoyed the most if they were playing a sport. All but one participant indicated that it was the "In the Middle of Easy and Hard" pile because these are activities that are "Just Right" or "Perfect" for them because they will not be boring like the too easy pile or frustrating like the too hard pile.

To provide further confirmation of these results, a latent partition analysis was conducted. This type of analysis uses a data matrix of probabilities (see Table 4.5) which has been generated from the number of times each item is grouped with every other item. Hierarchical cluster analysis was then used to sort the matrix into clusters of items that were sorted together by the participants. The items are said to be relevant and representative of the construct if the clusters match the categories from the instrument. As demonstrated in Figure 4.6, three distinct clusters emerged with the items clustering in the hypothesised category.

Individual Results: Sorting Techniques

A total of 14 of the 15 participants completed this portion of the study. As Table 4.6 demonstrates, participants tended to sort the items into three piles and interpreted the meaning of the items within the three piles to be: Hard, Middle, and Easy. Five of the participants in Table 4.6 sorted the items into piles and named those piles the same way in which they were intended. Two participants had all but one item placed in piles for which they were intended. In both cases, an item from the "Challenge Higher than Skill" category was placed in with items from either the "Challenge Equals Skill" items or the "Skill Higher than Challenge" items. Participant six decided to divide the items into four categories. All items from the "Challenge Equals Skills" subscale and the "Skill Higher than Challenge" subscale were grouped together. Items from the "Challenge Higher than Skill" subscale were grouped into two categories: Hard and Not too Sure.

Discussion

The results from this study provide evidence that the items are indeed relevant of children's perceptions and levels of optimal challenge (i.e., skill > challenge; challenge = skill; and, challenge > skill). Both the rank ordering technique and the sorting technique provided spatial solutions at the group and individual levels which indicated that participants rated items from the same subscales as similar to each other and different from items representing one of the two other subscales. The dimensionality interpretation also demonstrated a strong theoretical link to organismic theories of intrinsic motivation that support the importance of optimal challenge. However, the three dimensional solution did have a adequate goodness-of-fit indices. Future research may wish to examine the potential meaning of this third dimension using a theoretical interpretation. Finally, when using the sorting technique, the participants tended to label the items with descriptions such as "Hard", "In the Middle", and "Easy". These are the most common labels that children in previous research have attached to the three different levels of task difficulty (Mandigo, 2001a).

As indicated previously, validity is an on-going process. To ensure that scores do in fact provide accurate interpretations, it is important to ensure that the items are relevant to the construct. The procedures used in this study have many advantages in meeting these validation requirements. First, the procedures allow individuals to make direct judgements of psychological distances between stimuli (Kruskal & Wish, 1978). The participant ultimately decides how similar or dissimilar items are through direct comparison and therefore, the data is more likely to contain a relevant structure where experimenter contamination is minimised (Schiffman et al., 1981). This is especially important in instrument development for children. Often, it is assumed that children experience a psychological construct the same way as adults. Although this may or may not be true for some psychological constructs, it is vital that a researcher determine how children interpret the items that are meant to represent a psychological construct.

Secondly, by using the sorting technique, children ultimately decide if the researcher's interpretations of the items reflect their own interpretations. This is similar to establishing trustworthiness in qualitative research through member checking (Patton, 1990). The items for the self-report instrument used in this study were developed following interviews with children about their perceptions of optimal challenge (Mandigo, 2001a) and then revised following feedback from expert judges (Mandigo, 2001b). By using the sorting technique and having children label the piles they created provides evidence that the items initially generated from children's descriptions were relevant and representative of the construct.

The intention of this study is not to suggest that the method be used in the place of "traditional" validation procedures (e.g., factor analysis, internal consistency). Rather, it is intended to add another important source of evidence pertaining to the construct validity of test score interpretations. Further research is needed and warranted to provide construct validity evidence of the instrument's internal and external structures. For example, exploratory and confirmatory factor analyses could be used to provide internal construct validity evidence by examining the relationship among the items. As well, discriminant and convergent sources of validity that are theoretically aligned can help to provide evidence of an instrument's external construct validity. These various sources of traditional validity evidence combined with the one used in this study serve to strengthen the argument that interpretations generated from the items for a particular instrument provide accurate reflections of the construct for which one is interested in studying.

Finally, the combination of idiographic and nomothetic data sources used in this study serve to strengthen the interpretations at both the individual and group levels. Often, only aggregate data is presented and an understanding of the construct at the individual level (the level where the most behavioural change can be made) is lost. As a result, sources of measurement error can often be overlooked and not considered due to combining the data together (Bouffard, 1993). In the current study, the group solution for both procedures provided a clear and interpretable two dimensional solution. Although the majority of the individual solutions supported this interpretation, there were some individual differences. As such, future validation procedures for this instrument need to take into consideration potential dispositional differences that may affect the results. One potential way to examine this is to use the INDSCAL procedure that compares individual proximity matrices together. However, the ADSCAL individual solutions from this particular study will be a valuable source of information when conducting future validation studies should concerns regarding the use of certain items emerge. For example, when examining the factorial structure of the instrument in the future, concerns may arise that an item is cross-loading onto two different factors. Rather than simply exclude the item based on this statistical procedure, investigators can go back to either the individual sorting or rating results to identify whether there were concerns for that item at the individual level. Based on that information and the factor structure, a decision can then be made as to whether modify or delete the item.

Making decisions about the validity of an instrument needs to be based upon sound evidence. Without this evidence, a decision to add, delete, or retain an item is at best a guess and prone to unseen sources measurement error. Although the process may be long and meticulous, the journey is worthwhile to ensure that the interpretations that are made from the test scores do indeed reflect the construct for which they were intended.

Table 4.1.

Matrix of Mean Similarity Ratings Among the 18 Items When Using the Conditional Rank Ordering Method

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	.000																	
2	2.071	.000																
3	1.071	2.214	.000															
4	2.143	1.500	2.214	.000														
5	2.071	1.429	2.143	2.214	.000													
6	2.071	1.714	1.857	1.786	1.429	.000												
7	2.786	3.071	3.071	2.929	3.286	3.286	.000											
8	2.857	2.929	2.643	3.286	3.857	3.429	1.643	.000										
9	3.000	3.286	2.857	2.786	3.429	3.286	1.786	1.214	.000									
10	3.143	3.429	2.857	3.357	3.429	3.286	1.786	1.357	1.429	.000								
11	2.500	3.214	2.857	3.000	3.429	2.857	1.643	1.857	1.786	1.786	.000							
12	3.071	2.929	3.000	3.143	3.357	3.143	1.929	1.429	1.571	1.571	1.643	.000						
13	2.786	2.429	2.429	2.714	2.071	2.500	3.143	3.571	3.571	3.571	3.571	3.429	.000					
14	3.143	2.500	2.714	2.714	2.143	2.643	3.286	3.571	3.643	3.571	3.643	3.429	2.000	.000				
15	2.714	2.214	2.357	2.500	1.929	1.857	3.786	3.857	4.000	3.929	3.643	3.714	1.714	1.500	.000			
16	2.786	2.214	2.500	2.714	2.214	2.357	3.500	3.714	3.786	3.786	3.571	3.643	1.786	1.714	1.286	.000		
17	3.143	2.143	2.714	2.214	2.286	2.643	3.714	3.786	3.786	3.643	3.571	3.769	1.286	1.714	1.857	1.571	.000	
18	2.714	2.429	2.429	2.857	2.714	2.714	3.500	3.429	3.786	3.286	3.571	3.643	2.000	1.500	1.429	1.571	1.786	.000

Table 4.2

Stimulus Co-ordinates for Two Dimension Group Solution and Two Dimension Solution for Grades Two and Three: Rank Ordering Method

Item	Stimulus Co-ordinates					
	Entire Group Dimensions		Grade 2 Dimensions		Grade 3 Dimensions	
	I	II	I	II	I	II
1	0.0396	1.1716	-0.1713	1.0153	0.1468	1.2676
2	-0.5925	0.5762	-0.4975	0.9381	-0.6954	0.4261
3	-0.2490	0.7078	-0.3797	0.9025	-0.1928	0.5862
4	-0.3796	1.0056	-0.5464	0.8913	-0.3258	1.1541
5	-0.9906	0.5968	-1.0034	0.6910	-0.9944	0.4958
6	-0.6634	0.9148	-0.5972	1.0060	-0.7789	0.8386
7	1.5459	-0.3826	1.4869	-0.1181	1.5828	-0.3800
8	1.7504	-0.286	1.6747	-0.1120	1.7542	-0.3487
9	1.8193	0.1796	1.5678	0.5873	1.8833	-0.2330
10	1.7007	-0.4667	1.7075	-0.5544	1.6740	-0.2576
11	1.6026	0.2398	1.6030	-0.1004	1.5269	0.5077
12	1.6680	-0.3545	1.6454	-0.1871	1.6819	-0.3144
13	-1.0076	-0.7169	-0.8914	-0.9619	-1.0518	-0.5719
14	-1.0617	-0.9634	-0.6747	-1.3357	-1.2229	-0.7351
15	-1.4925	-0.1587	-1.4872	-0.4896	-1.3807	-0.1832
16	-1.2983	-0.5368	-1.1253	-0.5979	-1.3555	-0.6467
17	-1.3322	-0.6134	-1.6007	-0.3757	-1.0414	-0.7531
18	-1.0521	-0.9131	-0.7110	-1.1189	-1.2104	-0.8525

* Note: Dimension I = Skill: Challenge Ratio; Dimension II = Enjoyment

Table 4.3.

Goodness-of-Fit Values across Individuals

Participant	3 Dimensional		2 Dimensional		1 Dimensional	
	Stress	RSQ	Stress	RSQ	Stress	RSQ
1	.2031	.6482	.2996	.5200	.5647	.2218
2	.1512	.8593	.1904	.8412	.3454	.6901
4	.1414	.8854	.1927	.8564	.3056	.7997
5	.1572	.8246	.2144	.7848	.3531	.6732
6	.1665	.7758	.2196	.7600	.4000	.5485
7	.1910	.7313	.2428	.7200	.3744	.6140
9	.1559	.8560	.2305	.8111	.2831	.7912
10	.2130	.5025	.3057	.3975	.4963	.2693
11	.1534	.8298	.2144	.7800	.3824	.5789
12	.1779	.7165	.2547	.6505	.4388	.4640
13	.1737	.7923	.2355	.7424	.3269	.7124
14	.1842	.6888	.2533	.6367	.3822	.5422
15	.1881	.6995	.2735	.6544	.5224	.2713

Table 4.4

Stimulus Co-ordinates for Two Dimensional Group Solution for Sorting Technique

Item Number	Stimulus Co-ordinates	
	Dimension I	Dimension II
1	-.4058	1.3985
2	-.4971	1.3190
3	-.3841	1.4018
4	-.3177	1.3850
5	-.6776	1.1142
6	-.5526	1.2723
7	1.3508	-.3763
8	1.3508	-.3763
9	1.4534	-.2884
10	1.4550	-.2722
11	1.4238	-.1255
12	1.4232	-.3105
13	-.6636	-1.0920
14	-1.0105	-.9853
15	-.9907	-1.0072
16	-.9907	-1.0072
17	-.9946	-1.0359
18	-.9720	-1.0142

Table 4.5.

Probability Matrix for Sorting Technique

	OC1	OC2	OC3	OC4	OC5	OC6	CS7	CS8	CS9	CS10	CS11	CS12	SC13	SC14	SC15	SC16	SC17	SC18
OC1	--																	
OC2	71	--																
OC3	93	64	--															
OC4	79	71	86	--														
OC5	64	71	64	64	--													
OC6	71	79	71	71	93	--												
CS7	7	0	14	21	0	0	--											
CS8	7	0	14	21	0	0	100	--										
CS9	7	0	0	7	0	0	79	79	--									
CS10	0	0	7	14	0	0	93	93	86	--								
CS11	21	21	14	21	7	14	57	57	79	64	--							
CS12	0	0	7	14	0	0	79	79	86	86	64	--						
SC13	0	21	0	0	29	21	57	57	7	7	0	14	--					
SC14	7	14	7	7	36	29	0	0	0	0	0	7	86	--				
SC15	14	21	7	7	29	21	0	0	7	0	7	7	79	93	--			
SC16	14	21	7	7	29	21	0	0	7	0	7	7	79	93	100	--		
SC17	0	21	0	0	29	21	0	0	0	0	0	7	93	86	79	79	--	
SC18	7	14	14	14	29	21	7	7	0	7	0	14	79	93	93	93	79	--

Table 4.6.

Idiographic Profiles Based on Sorting Technique.

Participant	Pile 1	Pile 2	Pile 3	Pile 4	Pile 5	Pile 6	Pile 7
1	7,8,9,10,12,13 (Too Hard)	1,3 (Middle)	14,15,16,18 (Too easy)	13, 17 (Have more than you need)	11 (Not very good)	5,6 (Perfect)	2,4 (Challenging, but can do it)
2	7,8,9,10,11,12 (Too Hard)	1,2,3,4,5,6 (Middle)	13,14,15,16,17,18 (Easy)				
3	7,8,9,10,11 (Too Hard)	1,2,3,4,5,6 (In the Middle of Easy & Hard)	12,13,14,15,16,17, 18 (Easy)				
4	4,7,8,9,10,11,12 (Not so good; very, very bad)	1,3 (Medium)	2,5,6,13,14,15,16,17 ,18 (Really good at it)				
5	9,10,11,12 (Really Hard)	1,3,4,7,8 (Need more activity and in the middle of easy and hard; In the middle and not very hard)	2,5,6,13,14,15,16,17 ,18 (Simple and Really Easy)				
6	9,11,12 (Hard)	1,2,3,4,5,6 (Middle)	13,14,15,16,17,18 (Easy)	7,8,10 (Not too sure)			
8	7,8,9,10,11,12 (Little Too Hard)	1,2,3,4,5,6 (Sort of in the middle … not too easy, not too hard)	13,14,15,16,17,18 (Really, Really Easy)				

Participant	Pile 1	Pile 2	Pile 3	Pile 4	Pile 5	Pile 6	Pile 7
9	7,8,9,10,11,12 (Hard)	1,2,3,4,5,6 (Middle of easy and hard)	13,14,15,16,17,18 (Easy)				
10	3,4,7,8,10,12,18 (Didn't have enough … complaining didn't like it … too hard)	5,6,13,14,17 (Too much of something)	1,2,9,11,15,16 (Already done it before)				
11	7,8,9,10,11,12 (Hard & Difficult)	1,2,3,4,5,6, (In the middle of Hard & Easy)	13,14,15,16,17,18 (Really Easy)				
12	7,8,9,10,11,12 (Challenging)	2,13,17 (In between easy & hard)	1,3,4,5,6,14,15,16,1 8 (Easy)				
13	7,8,9,10,12 (Way too hard)	1,2,3,4,6,11 (Between Hard & Easy)	Everything is So Easy (5,13,14,15,16,17,1 8)				
14	7,8,9,10,12 (Hard)	1,2,3,4,5,6,11 (Easy)	13,14,15,16,17,18 (Very Easy)				
15	7,8,9,10,11,12 (Too Hard)	1,2,3,4,5,6 (Not too hard & not too easy)	13,14,15,16,17,18 (Over … same, over the challenge. So easy, it was boring)				

Figure Caption

Figure 4.1. Game Board.

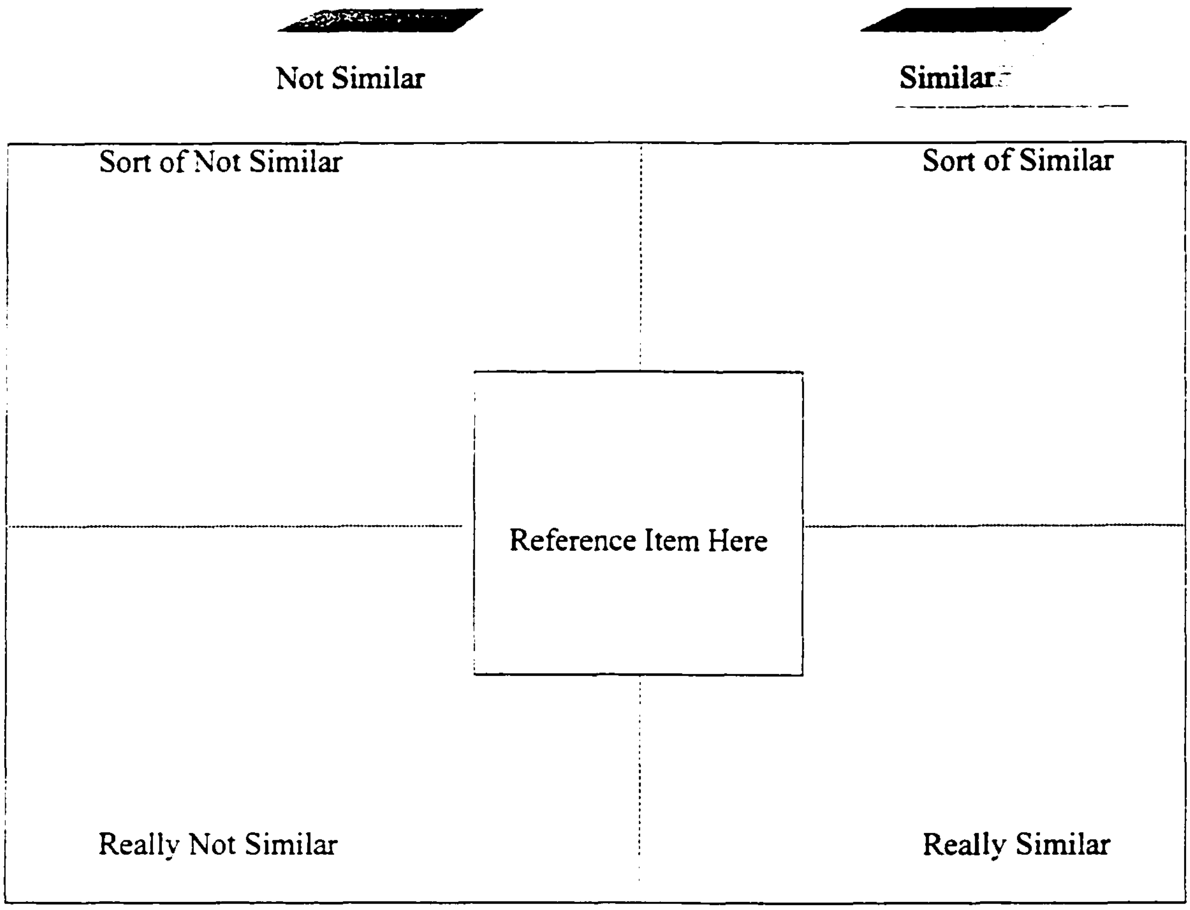

Not Similar
Similar
Sort of Not Similar
Sort of Similar
Reference Item Here
Really Not Similar
Really Similar

Figure Caption

Figure 4.2. Perceptions of different levels of optimal challenge items.

1. It was in the middle of easy and hard (S=C)
2. It was a good challenge that I could do (S=C)
3. It was not too easy and not too hard (S=C)
4. It was challenging, but I could do it well (S=C)
5. The challenge was perfect for my skill level (S=C)
6. My skills were equal to the challenge of the activity (S=C)
7. I needed more skill to do the activity (C > S)
8. It was too hard for me to do (C > S)
9. It was too challenging for my skill level (C > S)
10. I wanted to quit because it was too hard (C > S)
11. I was not very good at it (C > S)
12. The challenge was higher than what I could do (C > S)
13. I had more skill than I needed to do it (S > C)
14. It was so easy it was boring (S > C)
15. I could do it easily (S > C)
16. It was simple because I had done it lots of times before (S > C)
17. My skills were a lot higher than the challenge (S > C)
18. It was so easy, I did not want to do it any more (S > C)

SC = Skill equals Challenge; C > S = Challenge is Greater than Skill Level; S > C = Skill Level is Greater than Challenge

Figure Caption

Figure 4.3. Derived stimulus configuration for a two dimensional group solution from categorical rating technique.

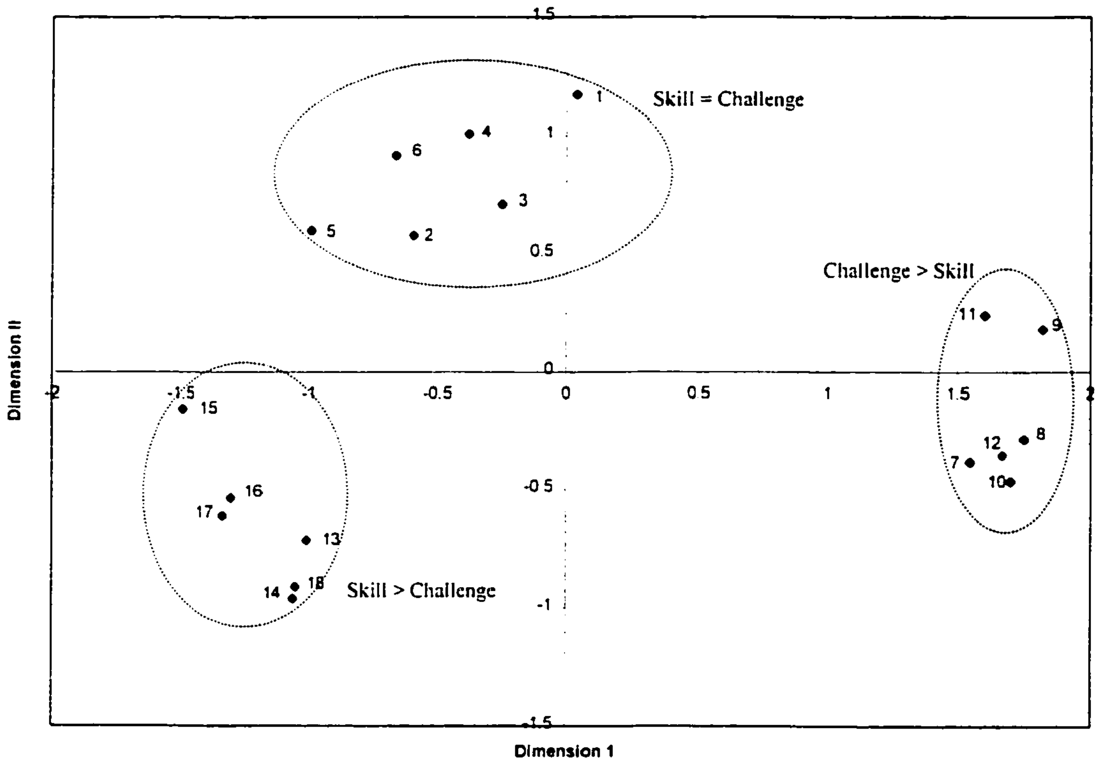

Skill = Challenge
Challenge > Skill
Skill > Challenge
Dimension II
Dimension 1
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

Figure Caption

Figure 4.4. Derived stimulus configuration for Participant 6.

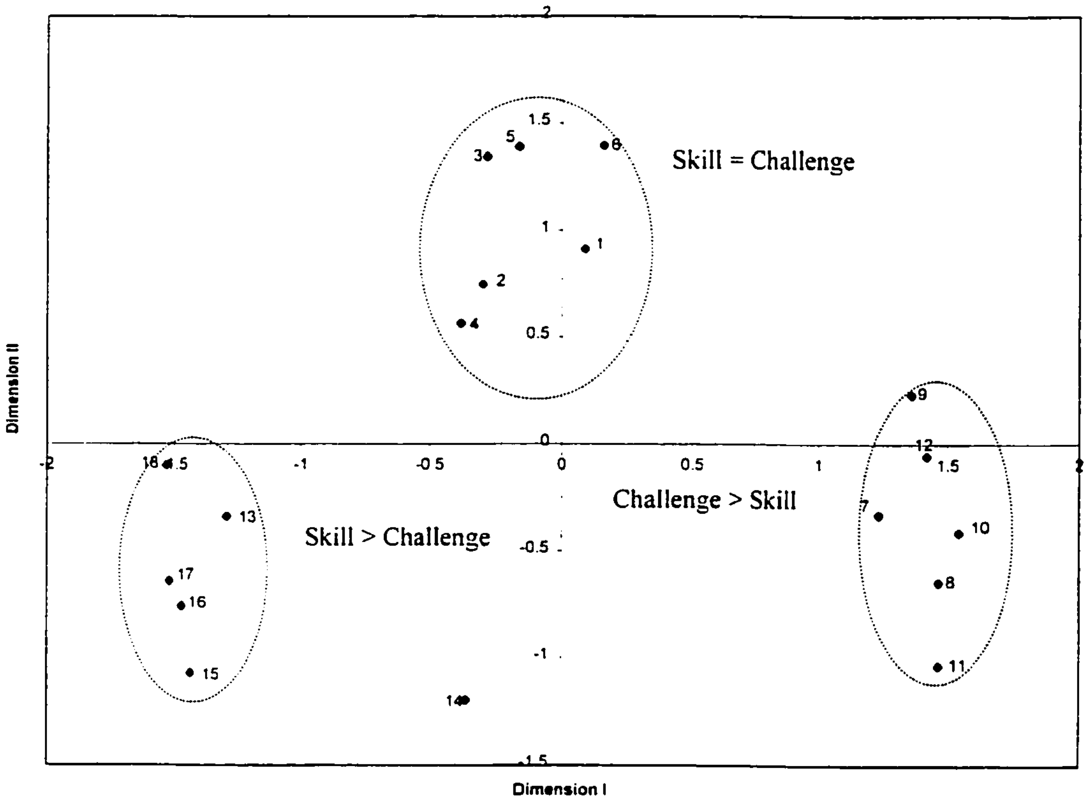

Skill = Challenge
Challenge > Skill
Skill > Challenge
Dimension II
Dimension I

Figure Caption

Figure 4.5. Group Solution: Sorting Technique

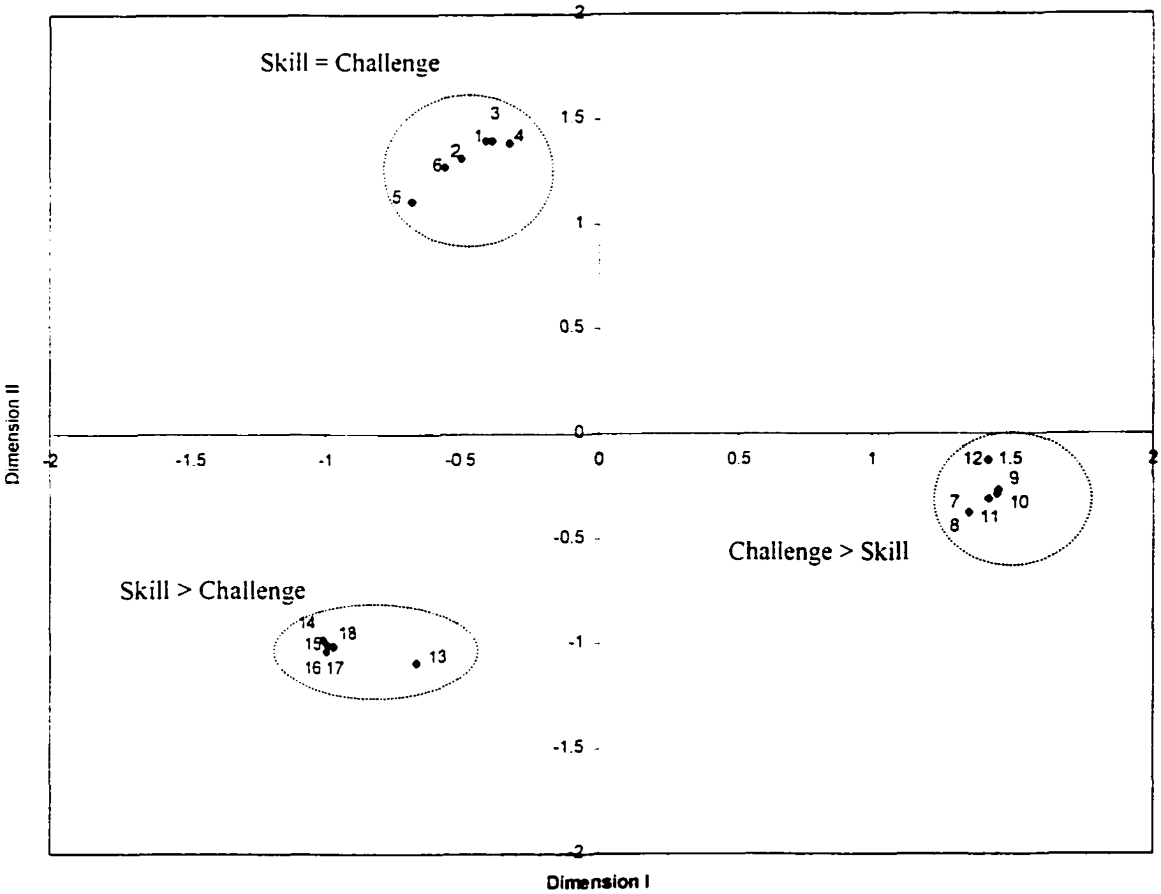
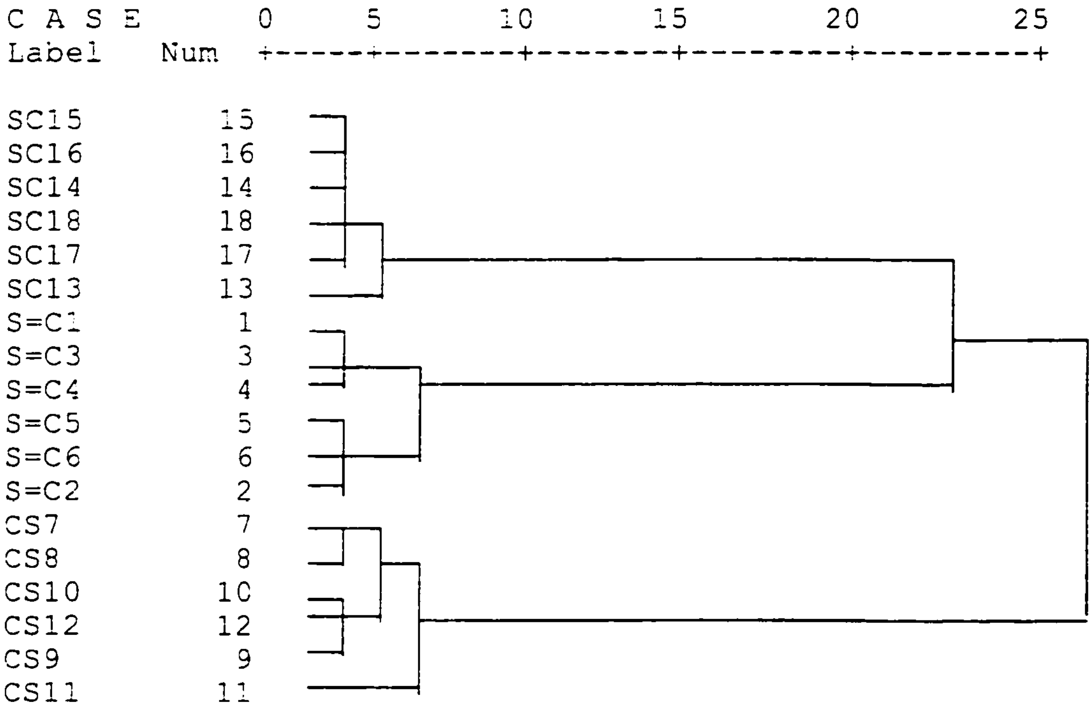

Skill = Challenge
Challenge > Skill
Skill > Challenge
Dimension I
Dimension II
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2

Figure Caption

Figure 4.6. Cluster analysis based upon proportion of times items were grouped together.


C A S E
Label Num
0 5 10 15 20 25
SC15 15
SC16 16
SC14 14
SC18 18
SC17 17
SC13 13
S=C1 1
S=C3 3
S=C4 4
S=C5 5
S=C6 6
S=C2 2
CS7 7
CS8 8
CS10 10
CS12 12
CS9 9
CS11 11

References

Biddle, S. J. H., & Chatzisarantis, N. (1999). Motivation for a physically active lifestyle through physical education. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand, & R. Seiler (Eds.), Psychology for physical educators (pp. 5 - 26). Champaign, IL: Human Kinetics.

Bouffard, M. (1993). The perils of averaging data in adapted physical activity research. Adapted Physical Activity Quarterly, 10, 371 - 391.

Brustad, R. J. (1998). Developmental considerations in sport and exercise psychology measurement. In J. Duda (Ed.), Advances in sport and exercise psychology measurement (pp. 461 - 470). Morgantown, WV: Fitness Information Technology.

Crain, W. (1992). Theories of development (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Philadelphia, PA: Harcourt Brace Jovanovich College.

Csikszentmihalyi, M. (1975). Beyond boredom and anxiety. San Francisco, CA: Jossey-Bass.

Davision, M. L. (1983). Multidimensional scaling. New York: John Wiley & Sons.

Deci, E. L. (1975). Intrinsic motivation. New York: Plenum.

Dunn, J. G. H. (1999). A theoretical framework for structuring the content of competitive worry in ice hockey. Journal of Sport and Exercise Psychology, 21, 259 - 279.

Flanery, R. C. (1990). Methodological and psychometric considerations in child reports. In A. M. La Greca (Ed.), Through the eyes of the child (pp. 57 - 108). Needham Heights, MA: Allyn & Bacon.

Gould, D., Feltz, D., & Weiss, M. (1985). Motives for participating in competitive youth swimming. International Journal of Sport Psychology, 16, 126 - 140.

Harter, S. (1974). Pleasure derived by children from cognitive challenge and mastery. Child Development, 45, 661 - 669

Harter, S. (1978). Effectance motivation reconsidered. Human Development, 21, 34 - 64.

Harter, S. (1985). Manual for the self-perception profile for children. Denver, CO: University of Denver Press.

Kruskal, J. B., & Wish, M. (1978). Multidimensional scaling. Beverly Hills, CA: Sage.

Mandigo, J. L., & Couture, R. T. (1996). An overview of the components of fun in physical education, organised sports and physical activity programs. Avante, 2, 56 - 72.

Mandigo, J. L. (2001a). Listening to the voices of children: Taking the first steps in the development of a self-report instrument of children's perceptions of optimal challenge. Unpublished manuscript. University of Alberta, Edmonton, Alberta, Canada.

Mandigo, J. L. (2001b). The importance of content-relevance evidence to support the construct validity of the children's perceptions of optimal challenge instrument. Unpublished manuscript. University of Alberta, Edmonton, Alberta, Canada.

Mautner, T. (Ed.), (1997). Dictionary of philosophy. London, UK: Penguin.

McCullagh, P., Matzkanin, K. T., Shaw, S. D. & Maldonado, M. (1993). Motivation for participation in physical activity: A comparison of parent-child perceived competencies and participation motives. Pediatric Exercise Science, 5, 224 - 233.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed.) (pp. 13 - 103). New York: American Council on Education.

Miller, P. H. (1993). Theories of developmental psychology (3rd ed.). New York: W.H. Freeman and Company.

Patton, M. Q. (1990). Qualitative evaluation methods. London, UK: Sage.

Petlichkoff, L. M. (1992). Youth participation and withdrawal. Is it simply a matter of fun? Pediatric Exercise Science, 4, 105 - 110.

Rowley, D. A. (1996). Physical activity through the eyes of six ten-year-old children: A case study. Unpublished dissertation. University of Alberta, Edmonton, Alberta, Canada.

Schiffman, S. S., Reynolds, M. L., Young, F. W. (1981). Introduction to multidimensional scaling. New York: Academic.

Stalans, L. J. (1997). Multidimensional scaling. In L. G. Grimm, & P. R. Yarnold (Eds.), Reading and understanding multivariate statistics (pp. 137 - 168). Washington, DC: American Psychological Association.

Weiss, M. R., & Bressan, E. S. (1985). Relating instructional theory to children's psychosocial development. Journal of Physical Education, Recreation and Dance, November/December, 34 - 36.

White, R. W. (1959). Motivation reconsidered: The concept of competence. Psychological Review, 66, 297 - 333.

CHAPTER 5

Study 4: Examining the psychometric properties of the Children's Perceptions of Optimal Challenge Instrument (CPOCI) in a Field Setting

In an era where concerns regarding the inactivity levels of children have been raised (e.g., Centre for Disease Control and Prevention, 1996; Craig, Russell, Cameron, Beaulieu, 1999), an increased importance has been placed on ways to increase children's intrinsic motivation to be active. Intrinsic motivation has been defined as occurring when someone is "... doing something because it is inherently interesting or enjoyable" (Ryan & Deci, 2000, p. 55) because he/she wants to feel competent and self-determined (Deci, 1975). According to three humanistic theories of motivation (Csikszentmihalyi, 1975; Deci & Ryan, 1985; Harter, 1978a), humans are highly motivated to take part in activities that are optimally challenging because of these potential outcomes.

Individuals are optimally challenged when they perceive the challenge(s) of the activity to be balanced with their abilities to do the task(s) (Csikszentmihalyi, 1990; Reeve, 1996; Weiss, 1986). The hypothesis put forth within these theories of motivation is that when participants feel optimally challenged during an activity, they are more likely to have a quality subjective experience and be intrinsically motivated to take part in the activity at that time and in the future due to enhanced perceptions of competence and enjoyment (Reeve, 1996). However, if exposed to continuous imbalances, participants may become frustrated or bored which may eventually lead to their withdrawal from the activity (Csikszentmihalyi, 1975; 1990). The importance of providing children with a positive physical activity experience such that they will want to remain physically active has been stressed by many as an important adherence factor (Centre for Disease Control and Prevention, 1997; Scanlan & Simons, 1992; Wankel, 1993). Providing for and creating optimally challenging experiences for children can be one of the ways to stimulate and maintain their motivation to continue participation.

Two of the first studies to point out the importance of optimal challenge were conducted by Harter (1974; 1978b) within the framework of Competence Motivation Theory (CMT). Harter's initial research suggested a curvilinear relationship with respect to challenge level and enjoyment. Harter also uncovered the negative effect of rewards (e.g., grades) on children's choice of difficulty level; however, the study did not take into consideration the initial ability level of the subjects. Danner and Lonky (1981) addressed this issue and found that children reported higher levels of interest and spent more time at stations that were either equal to or slightly higher than their skill level. An inverted-U relationship was also observed between interest ratings and perceived difficulty level. Tasks rated as either too easy or too hard were not rated as interesting as tasks that were matched according to the children's skill level. The effects of rewards (i.e., a good work certificate) were also investigated. The results demonstrated that when intrinsic motivation as measured by time on task and interest level for an optimally challenging activity was initially high, rewards decreased the children's motivation. When intrinsic motivation for a task perceived as being more difficult was initially low, rewards had little effect on motivation. The results supported the importance of providing children with choices and respecting their decisions when creating an optimally challenging environment and the negative effects that extrinsic rewards can have on participants' intrinsic motivation.

Rogers and Ponish (1987) examined the impact that giving children choice had on the level of task difficulty they chose. Their findings suggested that children were more likely to tolerate some level of self-imposed failure at more difficult tasks in order to attempt a higher level of challenge. However, in order to keep persisting, the children had to start experiencing some level of success, or else they would quit and/or move on to an easier task.

The previous studies all used objective measures of task difficulty (e.g., size of anagram, task difficulty) in relation to the participants' skill level. They did not directly assess the participants' perceptions of being optimally challenged based on perceived skill and perceived challenge. The theoretical framework for optimal challenge suggests that both the perceived challenge of the activity and the perceived ability level of the individual must be taken into consideration (Dobos, 1996). Because of the close theoretical link between the operational definition of flow and optimal challenge (i.e., both refer to the balance between perceptions of skill and challenge), several studies have used flow as a measure of optimal challenge. Experiencing flow (i.e., skills and challenge are equal) often results in a highly motivating experience (Csikszentmihalyi, 1975). Accordingly, the hypothesis is that the most enjoyable subjective experiences occur when individuals report a balance between perceived challenge and skill levels, while an unbalanced relationship between perceived skill and challenge often results in boredom or anxiety.

To date, there has been limited research conducted with children in physical activity environments that uses this ratio of perceived challenge and skill as an indication of the degree to which individuals are optimally challenged. Based upon participation in summer recreation programs, Mandigo and Couture (1996) and Mandigo, Thompson, and Couture (1998) found the flow state quadrant (i.e., higher than average skill and challenge) to be characterised by high quality of experience levels (i.e., high levels of fun, intrinsic motivation, positive affect, perceived success and low levels of state anxiety). However, high levels of subjective experience also characterised the boredom quadrant. Turner, Parkes, Cox, and Meyer (1995) found similar results with a group of eight children in a Grade 5 literacy class. In examining the characteristics of tasks that facilitated flow experiences, activities that provided students with a chance to modify and have some control were more likely to produce flow-like experiences (e.g., high levels of affect, potency/ activation, cognitive efficiency, and degree of engagement). Rowley (1996) reported that activities where skills and challenge were equal were viewed by all of the children as being the most fun. Alternatively, inappropriate challenges (i.e., too easy or too hard) were viewed quite negatively. Most participants indicated that they did not like activities that were not challenging or they were too challenging. As a result, many of the children in her study indicated that they actively sought out activities that challenged them in an appropriate manner. Sanders and Graham (1995) also concluded that children often resorted to "play" to get into situations where they could match their perceived abilities with the perceived challenge of the activities.

Despite these theoretical frameworks, no one theory adequately addresses the construct of optimal challenge. Many different theories have pointed to its importance, but they have not fully explored the conditions that create the psychological and behavioural outcomes of optimal challenge. Quite often, optimal challenge has been identified as facilitating intrinsic motivation (Ryan & Deci, 2000) and perceived competence (Harter, 1992), but the theories do not fully explain how to facilitate optimal challenge experiences.

One potential reason for this paucity of empirical evidence has been the lack of valid, developmentally appropriate instruments to tap into children's optimal challenge experiences during physical activities. In a review of instruments used within the sport psychology field, Vallerand and Fortier (1998) examined eight commonly used instruments to measure intrinsic motivation. Of these questionnaires, only the Motivational Orientations in Sports Scale (Weiss, Bredemeier, & Shewchuk, 1985) had a challenge subscale and this was intended to measure a trait. The Experience Sampling Form (ESF) commonly used to measure the flow states and optimal challenge experiences of individuals soon after participating in an activity (Csikszentmihalyi & Csikszentmihalyi, 1988) was not included in their review. The ESF provides a basic and unidimensional measure of optimal challenge by asking participants to report their perceived skill level and challenge of the activity soon after they have participated in an activity. However, some have questioned the validity of interpretations made from this instrument. For example, concerns over using only two items to measure optimal challenge (Ellis, 1987), the common procedure of standardising individual challenge and skill responses (Ellis, Voelkl, & Morris, 1994), the one dimensional perspective of skill and challenge (Voelkl & Ellis, 1998), and the validity of this measure with children (Mandigo & Thompson, 1998) have come into question.

Despite the theoretical importance of the construct of optimal challenge (e.g., Biddle & Chatzisarantis, 1999; Csikszentmihalyi, Rathunde, Whalen, 1997; Mandigo & Holt, 2000; Weiss & Bressan, 1985), the lack of a way to measure optimal challenge with children still exists. The following study is the fourth step in a series of studies intended to provide construct validation evidence of a newly developed instrument to tap into children's perceptions of various levels of optimal challenge. Test construction is an ongoing process that continually evaluates the validity of test scores (Golden, Sawicki, & Frazen, 1984; Linacre, 2000). Due to the high potential for measurement error when modifying existing adult instruments for children (Brustad, 1998), a series of studies were conducted. The first study involved extensive interviews with children about their perceptions of optimal challenge and the variables that surround it (Mandigo, 2001a). Generating items from interviews helps to enhance the understanding of the items for the participants and hence increases their relevance (Mahoney, Thombs, & Howe, 1995). After developing the items based upon the children's comments and descriptions, content relevance ratings by expert judges (Mandigo, 2001b) and direct similarity ratings between the items by children (Mandigo, 2001c) were conducted in studies two and three respectively. Based upon these initial three steps, the items were then placed into a questionnaire format deemed to be developmentally appropriate. The purpose of this study was to examine the psychometric properties of the CPOCI by examining its internal structure (e.g., factor structure and internal consistency) and how the subscales related to other measures associated with optimal challenge research (e.g., task difficulty, perceived skill, perceived challenge, perceived difficulty, and success). These steps were conducted to order to provide further construct validity evidence of the CPOCI. Collecting construct validation within an experimental activity setting is an important step in demonstrating the psychometric properties of an instrument (Cronbach, 1984; Messick, 1989).

Method

Participants

A total of 95 (57M; 38 F) participants between the ages of 7 - 12 years of age (M = 9.5 yrs; SD = 1.3) were recruited from a Summer Sport Camps hosted by a University in Southern Ontario, Canada. Participants were recruited through a covering letter sent home with each participant as well as a brief explanation of the study to parents and participants on the first day of the camp. Parents/ legal guardians and participants were asked to complete and submit an informed consent form should they wish to participate in the study. A member of the research team was available at the beginning of each day of the camp to collect completed informed consent forms and to answer any questions.

Procedure

Prior to each session, the lead researcher held discussions with the camp supervisor to determine the best time in the schedule for participants to take part in the study so that the amount of camp time participants missed was minimised. Participants were also consulted during the camp as to whether the time that had been predetermined was a good time for them to participate in the study. At that time, they were once again explained the purpose of the study, rights of a participant, and the general procedures of the study.

Each participant was taken to a "target" area, where he/she was asked to choose one activity from a list of five activities they wished to take part in. The choices for the activities were soccer, lacrosse, field hockey, underhand throwing, or overhand throwing. All of these activities were also part of the sports camp. Soccer was the most popular choice with 42% of participants choosing it for the study. This was followed by lacrosse (27%), overhand throwing (25%), field hockey (4%) and underhand throwing (1%). It was felt that providing participants with a choice of activities was important in order to avoid the negative impact of having participants take part in an activity that they did not enjoy. Due to the low number of participants choosing field hockey and underhand throwing, activities were combined into three categories for data analysis purposes: sending with hand (i.e., underhand and overhand throwing), sending with implement (i.e., field hockey and lacrosse) and sending with foot (i.e., soccer). A gender by age group breakdown by the number of participants participating in each category is provided in Table 5.1.

After participants chose an activity, they engaged in a brief warm-up that consisted of passing the object between partners. Participants then took part in the experimental task which consisted of trying to project an object (i.e., soccer ball, lacrosse ball, field hockey ball, softball, or beanbag) five times from each distance into the target (a hula hoop 60 cm in diameter). The experimental area (see Figure 5.1) consisted of a target and pylons marking distances from the target. Each pylon was set at a distance of 2.5 meters apart where the closest pylon was 2.5 meters and the furthest was 12.5 m from the target. These distances were chosen after carefully consulting various motor development references (e.g., Arnheim & Pestolesi, 1978; Corbin, 1980; Cratty, 1967; Hastad & Lacy, 1998). For example, DeOreo and Keogh (1980) reported a study where nine year old boys were able to successfully hit a 63.5 cm target in diameter from 7.62 m away at least two out of five throws. In addition, they reported a study that measured kicking accuracy into a 3m wide net from distances of 3m, 6m, and 9m away. Finally, many testing protocols for throwing accuracy have participants stand anywhere from 3 – 4.5 m away from a target for children between the ages of six to twelve (e.g., Arnheim & Pestolesi, 1978). The distances were chosen so that all participants would be able to get most of their shots in the target from the closest distance and where very few would be able to get any shots in from the furthest distance (i.e., 12.5 m). The order of the distances was randomised for each participant prior to the study in order to control for practice effects (see Table 5.2).

After completing a set of 5 "shots" at one distance, participants were asked to complete the Children's Perceptions of Optimal Challenge Instrument (CPOCI). Once they had finished completing the questionnaire, they then projected the object five times from a new distance. This continued until participants had five "shots" at each of the five distances and had completed the CPOCI for each distance. After completing the fifth distance, participants were asked to complete the CPOCI and the skill and challenge items from the Experience Sampling Form (Csikszentmihalyi & Larson, 1987) and Harter's (1974) perceived difficulty rating in addition to the CPOCI.

Once participants had completed all five distances, they were then asked to choose a distance that was "not too easy yet not too hard" for them. This sixth distance was recorded and they were given another five shots. They then completed the CPOCI one last time for that distance. Following this, they were asked to indicate the distance they felt was most enjoyable for them. Given that the time allotment was 30 minutes to participate in the study, only 65% (n = 62) of all participants entirely completed this sixth trial. This data was retained for evidence of predictive validity.

Instruments

Children's Perceptions of Optimal Challenge Instrument. The Children's Perceptions of Optimal Challenge Instrument (CPOCI) contains 3 subscales (skill = challenge, challenge greater than skill, skill greater than challenge) with 6 items per subscale (see Appendix A). The instrument is event-contingent and thus has been designed for participants to complete immediately following participation in a physical activity. Participants used an alternate choice item format by determining if the statement was true or false for them. Then, they decide whether it was really true/false or sort of true/false for them. Really True responses were coded as "4" while Really False responses were coded as "1". Participants used Crayola ® Stampers to stamp their responses to the questions. Items from the same subscale were summed together and an average for each subscale was calculated. The anchors used in the CPOCI are similar to those used by Harter (1985) and are assumed to be developmentally appropriate for children in this age range. Rust and Golombok (1989) recommended that four to seven options or descriptive anchors are needed in order to provide sufficient strength of response choices where participants have enough options to express themselves adequately. In order to control for social desirability, participants were asked to put down the answer that was right for him/her and to answer each question as quickly as possible (Rust & Golombok, 1989). As well, items from each subscale were presented every third question in order to control for acquiescence (Rust & Golombok, 1989). The advantage of using this type of questionnaire format is that the "true/ false" component is fast and easy to use and the descriptive anchors provide participants with an opportunity to express themselves more precisely by indicating the magnitude to which they agree or disagree with the statement (Rust & Golombok, 1989).

As indicated previously, the items for this questionnaire were developed through a series of studies that included interviews with children (Mandigo, 2001a), relevance ratings by expert judges of the items generated by study one (Mandigo, 2001b) and a review by 15 children to determine how similar and/or different participants thought the items left over from study two were to one another (Mandigo, 2001c). These items (see Figure 5.2) were then pilot tested with two children before being used in this study. The final version of the CPOCI used in this study had a Flesch-Kincaid Grade Level score of 3.1 as determined by a readability statistic in Microsoft Word. This score means that a third grader can understand the instrument. This was deemed appropriate for this sample as participants were in Grade 3 or above. This standard also meets previous recommendations that the readability of instruments for children should be at the third grade level (Stone & Lemanek, 1990).

Additional items. In addition to completing the CPOCI after the fifth trial, participants were asked to rate how challenging they felt the activity was and how much skill they felt they had. The items from the Experience Sampling Form (Csikszentmihalyi & Larson, 1987) were slightly modified and used for this portion of the study. Participants used a nine point Likert scale to indicate the degree to which they felt their skills were to do the activity and how challenging the activity was (1 = low; 9 = high). Participants were also asked to use a five point Likert scale to indicate how difficult they felt the activity was. The same Likert scale used by Harter (1974) was used for this item (1 = Very Easy; 2 = Easy; 3 = Medium; 4 = Hard; 5 = Very Hard).

Results

Internal Structure

The internal structure (i.e., factor structure and internal consistency) of the subscales across the various trials was examined initially to ensure the items from the same subscale did "cluster" together. Table 5.3 presents the means and standard deviations of each subscale across the six trials.

Principal component factor analysis. To explore the factorial structure of the three subscales, Exploratory Factor Analysis using a principal components analysis with a varimax rotation was used. Using Crocker and Algina's (1986) recommendations for accepting factors with eigenvalues greater than one, a four factor solution was recommended. However, an examination of the scree plots suggested a three factor solution. A forced three-factor solution using a principal component analysis with a varimax rotation was then conducted and compared to the four factor solution generated by the exploratory factor analysis. Each of the three factors met the eigenvalue requirements and had a better simple structure (i.e., fewer cross loadings) than the four factor solution generated by the initial exploratory factor analysis. Therefore, due to the interpretation from the scree plots for both factor analyses and the better simple structure in forced three factor analysis, it was decided that the three factor solution was the most appropriate and more interpretable theoretically. To make comparisons across each of the four trials, forced three-factor principal components analyses were used to determine whether the items loaded on the appropriate subscales. Based on Brynt and Yarnold's

(1997) requirements of five times the number of participants as items for factor analyses, trial six (i.e., the choice trial) was not factor analysed because only 62 participants completed this "bonus round" due to time restrictions. An analysis of the scree plots and a review of the eigenvalues demonstrated that the three-factor solution was appropriate for the remaining four trials. As indicated in Table 5.4, all of the items fit the hypothesised factor structure across all five trials with some cross-loadings based upon a minimum factor loading standard of .30 (Crocker & Algina, 1986). However, the majority of these cross-loadings were in the negative direction for the skill greater than challenge and challenge greater than skill subscales. This suggests that some items were negatively correlated to items in the opposite subscale. This is not surprising given that these two subscales are the opposite of each other. For example, item 3 on the skill greater than challenge subscale [I could do it really easily] loaded negatively onto the challenge greater than skill factor four times. This suggests that this item has a negative relationship to the challenge greater than skill subscale. The amount of explained variance for all five factor analyses ranged from 51% for the first trial to 65% for the trial five.

Internal consistency. Cronbach alpha coefficients were used to determine the internal consistency of the three subscales from the CPOCI. Minimal acceptable standards for alpha coefficients are usually between .70 (DeVillis, 1991) to .75 (Mahoney, Thombs, & Howe, 1995). As indicated in Table 5.5, only one subscale across all five trials was below this minimal standard. All other subscales across the five trials had alpha coefficients above .75. Cronbach alpha was also computed for age groups (see Table 5.6) and gender (see Table 5.7). As Table 5.6 indicates, those in the younger age group (i.e., ages 7 - 9 years) had lower alpha coefficients at the initial trials compared to the older participants (i.e., ages 10 - 12 years) who consistently had acceptable levels across all trials. Both genders reported acceptable levels (or near acceptable levels) across all trials.

Corrected item-total correlation. The corrected item-total correlation was also used to determine the correlation between an item's score and the scale score. To avoid artificially inflating the correlation, the item score was not included in the total subscale score when conducting the correlation with the particular item. This correlation is useful in determining the degree to which an item represents what the test measures as a whole and has been recommended to be in the range of .40 - .60 (Golden, Sawicki, & Frazen, 1984). This range allows for sufficient redundancy among items while at the same time, ensures that the items are not exactly the same. The correlations presented in Table 5.8 demonstrate that most of the correlations (91%) were between .40 and .80. Although many items exceeded the .60 suggestion by Golden et al. (1984), it does demonstrate some level of redundancy that is an important consideration in questionnaire construction as indicated by DeVillis (1991).

Construct Validity Evidence

To ensure that the distances were progressively more difficult from each other, the within-subject factor of successful hits (5 distances) and the between-subject factors of gender (2), age group (2) and type of activity (3) were analysed using a repeated multivariate analysis of variance (MANOVA) in SPSS 10.0. The age group variable consisted of two levels: 7 – 9 and 10 – 12 years of age. The type of activity variable was broken into three levels: project with implement, project with hand, and project with feet. This was done to ensure that there were no empty cells within the analysis. Using Wilks' criteria, a significant multivariate effect was uncovered for the within-subjects effect of distance ($\Lambda_{4,79} = .232$; $p < .001$; $eta^2 = .77$) and for the distance by type of activity interaction ($\Lambda_{8,158} = .813$; $p < .05$; $eta^2 = .10$). Figure 5.3 provides a representation of the differences in successful attempts across each distance for the entire group and for the different type of activities. Follow-up pairwise comparisons using the Bonferroni correction factor were used to determine where the differences were. This correction factor uses Student's t statistic to protect against Type one errors when making multiple comparisons. The results of this comparison indicated only distance four and five did not differ significantly ($p < .05$) on successful attempts for the entire group. Based upon the significant interaction effect that was found, the group means for number of times the target was hit for each type of activity (see Figure 5.3) were compared using pairwise comparisons with the means adjusted again for multiple comparisons. These results showed that those who chose activities that sent the object with an implement (i.e., lacrosse and field hockey) hit the target significantly ($p < .05$) fewer times across the five distances than those who chose to project the object with their hand (i.e., underhand or overhand throwing) or with their feet (i.e., soccer).

Due to this finding, it was necessary to test for potential group differences on subscale means. A doubly-multivariate analysis of variance as described by Tabachnick and Fidell (2001) was performed on the three optimal challenge subscale scores and three between-subject factors. The within-subject variables treated multivarietly were the mean subscale scores (three subscales measured repeatedly). The between-subject factors of gender (2), age group (2) and type of activity (3) were included in the GLM to test the hypothesis of potential group differences. As indicated in Table 5.9, the only within-subject multivariate effect was for the subscale means across the various distances. The only between-subject multivariate effects were for the age group ($\Lambda_{3,79} = .852$; $p < .01$) and type of activity ($\Lambda_{6,158} = .827$; $p < .05$) variables. There were no significant within or between multivariate interactions.

Based upon the significant multivariate within-subject effect for subscale scores across the five distances, means for each subscale were compared to each other across the five distances (see Figure 5.4). Using Student's t statistic and correcting for multiple comparisons using a Bonferroni correction factor, there were several significant ($p < .05$) differences. First, for the skill equals challenge subscale, the mean for distance three was significantly higher than distance one and five. For the skill greater than challenge subscale, distance one had a greater mean than the other four distances, distance two had a higher mean compared to distance four and five, and distance three had a higher mean compared to distance five. Finally, for the challenge greater than skill subscale, only distance three and four did not have significantly different means. All other means increased significantly across the distances.

Based upon the significant between-subject multivariate effect for age group and type of activity, univariate between-subject effects were examined for the two variables. Subsequent between-subject effects indicated two significant main effects. For the type of activity, there was a significant main effect for the challenge greater than skill subscale ($F_{2,81} = 6.50$; $p < .01$; $eta^2 = .138$). Follow-up pairwise comparisons based on Student's t statistic and using the Bonferroni correction factor found that those who sent the object with an implement reported higher means on this subscale than those who sent the object away with their feet. The other main effect was for age group. More specifically, there were between-subject effects for the skill equals challenge subscale ($\underline{F}_{1,81} = 4.06$; $\underline{p} < .05$; $eta^2 = .048$) and the skill greater than challenge subscale ($\underline{F}_{1,81} = 7.32$; $\underline{p} < .01$; $eta^2 = .083$). Follow-up pairwise comparisons using the Bonferroni correction factor indicated that younger participants (i.e., 7 to 9 years) reported a higher mean for the skill equals challenge subscale and a lower mean for the skill greater than challenge subscale across all five distances than the older participants (i.e., 10 to twelve years).

Given the perils of averaging data (Bouffard, 1993; Livingston & Mandigo, 1999), individual profiles were developed by graphing the means for each subscale across each distance. The structure of each graph (i.e., the formation of the lines across the X-axis) was examined through visual inspection and compared to other individual profile graphs. Graphs that had similar structures were then grouped together and compared to the group solution (see Figure 5.4). One quarter (25%) of the individual profile plots matched the same structure as group profile (i.e., inverted U shape for Skill = Challenge; descending line for Skill > Challenge; and, ascending line for Challenge > Skill subscales across distances). These profiles represent the hypothesised means for each of the three subscales. Another 28% of the profiles closely resembled the hypothesised means across the distances, except for one or two of the distance subscale scores falling outside of the predicted pattern. A few individual profiles (3%) had two out of the three subscale scores matching the hypothesised structure across the various distances. Approximately 4% of profile plots only had the skill equals challenge subscale matching the hypothesised structure while 1% of the plots only matched the skill greater than challenge structure and 1% of plots only matched the challenge greater than skill structure. About 6% of the profile plots only fit the trial 6 prediction of high skill equals challenge score and moderate scores on the other two subscales. Five percent of the plots were grouped into a category where the subscales fit theoretical predictions, however, they scored high on the skill equals challenge subscale at distance one and this score went down across trials. Alternatively, 7% of the plots were in the opposite direction with the skill equals challenge subscale being low to begin with and increasing across distance. Approximately 14% of the plots were grouped into a category called "parallel". That is to say, the scores from the three subscales tended to parallel each other across the various distances. This suggested that these participants may have rushed through the questionnaire in order to do the activity, did not understand the items, and/or did not make a distinction between distances. Finally, approximately 4% of the profile plots could not be placed into a category because there was no evident pattern to the plots.

On the fifth trial, participants completed the CPOCI and rated their perceptions of skill, challenge, and task difficulty. These additional items and the distance and success participants had during this trial were correlated against the three subscales to provide concurrent validity evidence (see Table 5.10). As hypothesised, the skill equals challenge subscale had moderate to low correlations with these variables. This is not surprising given that when an activity is not too challenging or where skill level is not too high that participants should report higher means on the skill equals challenge subscale. Contrary to Harter's (1978b) Competence Motivation Theory, the skill equals challenge subscale was negatively correlated to the number of times participants hit the target. For the other two subscales, the correlations were highly significant and in the expected direction. For example, perceived challenge was correlated positively to the challenge greater than skill subscale and correlated negatively to the skill greater than challenge subscale.

Due to the highly motivating experience of activities where participants perceived their skills and the challenge to be balanced (Csikszentmihalyi, 1990; Deci & Ryan, 1985; Harter, 1978a) it was hypothesised that when given a choice, participants would choose the distance where they scored the highest on the skill equals challenge subscale. To test this relationship, a repeated measures MANOVA as outlined in the General Linear Model was used to examine the within-subject effect of subscale scores across the five distances plus the distance that was chosen on the last trial (i.e., Distance = 6). A significant multivariate effect was found for the within-subject variable($\Lambda_{15,47} = .193$; $p < .001$; $eta^2 = .81$). Figure 5.4 demonstrates that participants scored highest on the skill equals challenge subscale when they chose the distance that was not too easy, not too hard for the last trial and scored in the mid-range on the other two subscales. Subsequent pairwise comparisons using the Bonferroni correction factor demonstrated that for the skill equals challenge, means were significantly higher than the other distances. In addition, participants reported significantly ($p < .05$) lower skill greater than challenge scores on the choice trial when compared to distance one and higher scores compared to distance five. Finally, for the challenge greater than skill subscale, participants reported significantly higher means for the choice trial when compared to distance one and significantly lower means when compared to distance four and five.

At a more individual level, 31% of participants chose the same distance for the sixth trial as the one they scored the highest on the skill = challenge subscale while 28% of participants reported that their most enjoyable distance was the one in which they scored the highest on the skill equals challenge subscale. Approximately 32% of participants chose the same distance as the one that was not too easy and not too hard as the one they chose that was the most enjoyable for them. In addition, 55% of participants chose distance three as the one that was not too easy, yet not too hard.

Discussion

Overall, the psychometric properties of the CPOCI were supported. The factor structure and internal consistency of the items was supported across most of the experimental trials. Individual items were also consistently correlated with the total subscale score. The validity of the interpretations made from the CPOCI was also supported through various measures. First, the various mean subscale scores fit their hypothesised relationship with the distance from which they were aiming (see Figure 5.4). Based upon the distance from the target, there was a curvilinear pattern for the skill equals challenge subscale means, a negative relationship with the skill greater than challenge subscale means, and a positive relationship with the challenge greater than skill subscale means. As well, when asked to stand at a distance that was not too hard and not too easy, participants scored high on the skill equals challenge subscale and moderate on the other two subscales. Finally, concurrent validity was demonstrated through anticipated correlations between subscale scores and other measures associated with optimal challenge (i.e., perceived skill, challenge, and task difficulty).

Given the fact that validity is a never-ending process (Messick, 1989), further research is needed to understand some of the differences that occurred in this study. First, the psychometric properties of the CPOCI seemed to improve after the first trial. That is to say, the factor loadings and the alpha coefficient values were higher and more consistent after the first trial. When breaking this relationship down by gender and age group, it appeared that the older participants (i.e., 10 - 12 years) initially understood the items better than the younger age group (i.e., 7 - 9 years). This may point to a developmental issue in relation to procedures when using self-report instruments for younger children rather than being an issue with the instrument itself. Perhaps, they need to receive a practice trial before completing the questionnaire as part of the study. Future research may be warranted as a result of these findings.

Secondly, despite the group plot meeting the hypothesised predicted structure, only 53% of the individual plots either matched or closely matched the group graph of distance by subscale means. This suggests that potential individual differences may be masked as a result of only using the aggregate (Bouffard, 1993). For example, in 17% of the plots, participants either initially scored high or low on the skill equals challenge subscale. One potential hypothesis to explain this result is that experienced participants may score high on the skill equals challenge subscale at higher levels of task complexity because they have mastered the skills for the task. Alternatively, lesser experienced and/or skilled participants may score high at lower levels of task complexity because they find "easier tasks" harder when first trying a novel task. Given this finding, future research with the CPOCI should consider obtaining information on participants' prior experience for the task at hand and how participants perhaps transfer skills from familiar sports to ones they are less familiar with.

Although the individual profiles were valuable in suggesting potential individual differences at the idiographic level, the method used to arrive at this conclusion may be cause for concern. In the current study, the lead researcher sorted the graphs into piles based on the similarity between the structure. Future research may wish to ask a second person to sort the graphs into piles and then examine the degree of agreement between the researchers. Future researchers may also wish to use the individual profiles and go back to participants in the study to further explore the reasons why they rated certain distances the way they did or how they interpreted the questions on the questionnaire.

It was also hypothesised that on trial six, participants would choose the distance where they scored the highest on the skill equals challenge subscale. However, only 31% of participants did this. This could be attributed to a couple of reasons. First, there may have been some participants who felt they improved as a result of having taken 25 shots at a target and hence, would have felt a more difficult distance fit this criteria. For others, they may have felt they were being evaluated despite every attempt to make them feel like it was not a test. For these individuals, they may have chosen an easier task. There were also others who may have been disappointed with their success rate. Many of the lacrosse participants expressed some frustration with not being able to hit the target and hence when given a choice, may have chosen an easier distance in order to experience success. Future studies that use this method may wish to ask participants why they chose the distance that they did. As well, when exploring the predictive validity of the CPOCI, one may wish to collect data over a period of time and across various types of activities in order to determine it's longitudinal impact on motivation and adherence.

Although the group profile plot demonstrated that the subscale means across the various distances fit the hypothesised structure, there were some significant group differences in the way the subscales were answered as a result of type of activity. Although each activity had the same subscale mean by distance structure when graphed, the peak for the skill equals challenge subscale occurred at different distances depending on the type of activity. For example, participants doing activities of sending away with foot reported the highest skill equals challenge scores at distance four, while those who used their hands or an implement to send away the object peaked at distance three. These results demonstrate the need to examine participants' perceptions of sport specific competence and it's relationship to optimal challenge. Based upon Harter's (1978a) competence motivation theory, it would be hypothesised that those with lower levels of perceived competence for a particular task would report being optimally challenged at lower levels of task complexity than those with higher levels of perceived competence. In the current study, many participants who chose lacrosse because it was a novel task and thus lacked the necessary skills to successfully put the ball into the target. Given that perceived competence is commonly correlated positively with prior experience (Feltz, Gould, & Horn, 1982; Roberts, Kleiber, & Duda, 1981), it is not surprising that these participants reported being optimally challenged at closer distances. For an activity such as soccer, however, more may opportunities exist for participants to play the sport at school and in the community and as such, likely have more experience and higher levels of competence regarding their soccer skills. As such, they tended to report higher levels of skill equals challenge scores at distances further away. Perhaps if the instrument had been used with children with similar experiences and/or who were engaging in the same task/activity, then there may have been more of the individual profiles matching hypothesised patterns.

There was also a significant difference between age groups for the skill greater than challenge subscale. Again, the shape of the graph for both age groups was similar (i.e., the means decreased across distance), but the older age group tended to report higher means than the younger age group suggesting that they found the activities easier. This seems reasonable given that the activities were skill oriented and likely easier for the older age group. Future research may wish to examine the different levels of optimal challenge in more complex game scenarios where participants are required to use tactics and strategies in conjunction with their skills.

Finally, some of the individual items did cross-load onto other factors across the five separate factor analyses and the sample size was slightly under the recommended minimum of 100 (Bryant & Yarnold, 1997). Further investigation is warranted to explore the factor structure of the CPOCI with a larger sample and to determine how participants were interpreting these items. For example, the challenge greater than skill item of "I wanted to quit because it was too hard" cross loaded during two trials onto the skill greater than challenge subscale. As well, the skill greater than challenge item of "I could do it really easily" loaded four times onto the challenge greater than skill subscale but was negative. Although the factor loadings were higher on three of the four subscale factors for which they were intended, further investigation of potential alternate interpretations of these items is warranted to determine if children have more than one interpretation of the items (Crocker & Algina, 1986). In addition, because the forced three factor solution had the best simple structure when compared to the four factor solution for trial 1, all of the remaining factor analyses had a forced three factor solution for sake of comparison. Future research will need to determine whether a three factor solution is the best solution for the CPOCI or whether a different factor structure is the best representation. For example, on Trial 3, there were many cross-loadings in the negative direction for the skill greater than challenge and challenge greater than skill subscales. This suggests that for this trial, a two factor solution (e.g., skill = challenge and skill ≠ challenge) may have been the best solution. Caution must then be taken, however, to ensure that the interpretations made from the factor structure are considered in conjunction with the theoretical framework from which the CPOCI has been developed.

In lieu of the current findings, it is recommended that the CPOCI continue to be used as it appears to be beneficial for group analysis. However, more research should be done to further validate the instrument's use for individual analysis so that the amount of measurement error can be reduced even further. For example, a child's goal orientation has been suggested to have an impact on his/her preference for optimally challenging activities. Sarrazin and Famose (1999) reported that participants' choice of a climbing wall was influenced by their goal orientations and perceived ability. Participants who were ego-oriented but had low levels of perceived climbing ability, tended to choose easy courses. About 25% of these participants, however, tended to choose the very difficult course so that they would fail on purpose and hence, attribute their failure to the difficulty of the wall rather than their skill. For those who were ego-oriented yet had high levels of perceived climbing ability, they were more likely to choose the moderate over the difficult course. Participants, who were high on task-orientation, yet low on perceived ability tended to choose the moderate course while those high on both task-orientation and perceived climbing ability exclusively chose the difficult courses. None of the task-oriented participants chose an easy course with a high probability of success.

Boggiano, Main, and Katz (1988) also investigated the impact that various dispositional and situational variables have on children's preference for challenging tasks. In their first study, they found that children's self-reported perceptions of academic competence and personal control were positively related to intrinsic interest in schoolwork and preference for challenging school activities. In the second study, the authors reported that children with high levels of academic competence and personal control were more likely to report higher preferences for challenging activities when placed in evaluative, and controlling conditions than those with low levels of academic competence and personal control. However, when the controlling condition did not exist, there were no differences between those with high and low levels of academic competence and personal control.

The CPOCI can be used as a valuable pedagogical tool for researchers to help discover the environmental conditions that should be in place in order to facilitate optimal challenge. For example, using experimental and factorial designs, future research can examine the impact that certain situational constraints (e.g., the use of rewards, various teaching styles) have on different dispositional variables (e.g., goal orientation, preference for challenging activities) when trying to create optimally challenging environments. In doing so, activity practitioners can be better informed as to various ways to facilitate optimal challenge during their classes.

Although this study follows three previous studies, it is just the beginning. Validation is a never-ending process. We can never say that something is truly valid because self-report measures will always contain some measurement error (Crocker & Algina, 1986). However, the more known sources of measurement error are controlled, the closer the observed score will be to the real score. As this study demonstrated supportive validation evidence, no instrument is perfect. The CPOCI, however, is the best available tool at the moment to capture children's level of optimal challenge and has demonstrated usefulness with children as participants. It has the potential of taking researchers one step closer to uncovering how to optimally challenge children in their physical activity settings.

Table 5.1.

Number of Participants by Age Group, Gender, and Activity Chosen

	Gender				Total
	Male		Female		
Activity Category	7 – 9 Yrs	10 – 12 Yrs	7 – 9 Yrs	10 – 12 Yrs	
Send with Foot	8	8	13	11	40
Send with Hand	3	12	3	7	25
Send with Implement	24	2	2	2	30
Total	35	22	18	20	95

Table 5.2.

Randomisation of Trials: Percentage of Participants by Trial Distances

	Trial				
Distance	1	2	3	4	5
1	17.9	25.3	17.9	16.8	22.1
2	21.1	17.9	21.1	17.9	22.1
3	27.4	17.9	14.7	23.2	15.8
4	16.8	21.1	21.1	20.0	20.0
5	16.8	17.9	24.2	21.1	18.9

Table 5.3.

Means and Standard Deviations of Subscales Across Trial

	Trial									
	1		2		3		4		5	
Subscale	M	SD	M	SD	M	SD	M	SD	M	SD
Skill = Challenge	3.03	.61	2.83	.82	2.71	.83	2.78	.87	2.67	.91
Skill > Challenge	1.93	.65	2.04	.75	1.92	.73	2.00	.77	2.22	.85
Challenge > Skill	1.93	.71	1.85	.72	2.03	.80	1.89	.74	1.88	.78

Table 5.4.

Forced Three-Factor Principal Component Analyses Across all Five Trials

	Trial 1 Factors			Trial 2 Factors			Trial 3 Factors			Trial 4 Factors			Trial 5 Factors		
Item	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
S=C1			.546	.382	.497			.731		.782			.760		
S=C2			.615		.687			.707		.738			.716		
S=C3			.672		.756			.667		.800			.820		
S=C4			.784		.840			.791		.732			.797		
S=C5			.602		.808			.770		.852			.853		
S=C6			.381		.813			.822		.872			.771		
CS1	-.364	.494		.766			.856				.842			-.327	.719
CS2		.828		.766			.853				.801			-.319	.793
CS3		.794		.692			.798				.762				.789
CS4	.398	.553		.640			.524		.438		.581	.392			.765
CS5		.645		.710		-.319	.825				.733			-.390	.715
CS6		.716		.668			.724				.766				.793
SC1	.641					.797		-.358	.579			.761		.788	
SC2	.750					.777	-.498		.636			.754		.717	
SC3	.776			-.541		.638	-.728		.437		-.476	.706		.730	-.430
SC4	.748					.656	-.532		.535			.671		.676	
SC5	.724					.677	-.532		.584			.779		.755	
SC6	.685					.676			.755			.796		.827	

Table 5.5.

Cronbach Alpha Coefficients Across Trials

	Trial				
Subscale	1	2	3	4	5
Skill = Challenge	.6822	.8324	.8463	.8836	.8842
Challenge > Skill	.7854	.8269	.8727	.8615	.8767
Skill > Challenge	.8262	.8413	.8506	.8672	.8654

Table 5.6.

Cronbach Alpha Coefficient by Age Group Across Trials

	Trial				
Subscale	1	2	3	4	5
Skill = Challenge					
7 - 9 Yrs	.51	.76	.80	.87	.87
10 - 12 Yrs	.81	.90	.90	.90	.90
Challenge > Skill					
7 - 9 Yrs	.69	.79	.85	.84	.83
10 - 12 Yrs	.84	.85	.88	.84	.93
Skill > Challenge					
7 - 9 Yrs	.79	.79	.85	.88	.82
10 - 12 Yrs	.85	.88	.84	.84	.91

Table 5.7.

Cronbach Alpha Coefficient by Gender Across Trials

Subscale	Trial				
	1	2	3	4	5
Skill = Challenge					
Male	.63	.84	.86	.89	.88
Female	.75	.81	.83	.87	.88
Challenge > Skill					
Male	.79	.85	.87	.84	.82
Female	.78	.79	.88	.88	.93
Skill > Challenge					
Male	.81	.85	.86	.85	.86
Female	.86	.84	.83	.90	.88

Table 5.8.

Corrected Item Total Correlations

	Trial				
Item	1	2	3	4	5
S=C1	.3796	.4267	.5754	.6585	.6745
S=C2	.3874	.5528	.5839	.6119	.6139
S=C3	.4498	.6271	.5723	.6927	.7580
S=C4	.6277	.7277	.7236	.6430	.7068
S=C5	.4202	.6737	.6542	.7768	.7475
S=C6	.2287	.6347	.6643	.7939	.6764
SC1	.5061	.6341	.5629	.5748	.6164
SC2	.6319	.6656	.7380	.7555	.6340
SC3	.6594	.6842	.6945	.7304	.7062
SC4	.6743	.6031	.7179	.6323	.5940
SC5	.6339	.6737	.7616	.7402	.7278
SC6	.4991	.4663	.3247	.5905	.7053
CS1	.4543	.7255	.7979	.8108	.7087
CS2	.7239	.7108	.7553	.7139	.6873
CS3	.6201	.5978	.7264	.6812	.6801
CS4	.2821	.3965	.3719	.3681	.5486
CS5	.5212	.6198	.7431	.6524	.6995
CS6	.6508	.5448	.6474	.7082	.7856

Table 5.9.

5 (Distance) x 2 (Gender) x 2 (Age Group) x 3 (Activity) Doubly Multivariate Repeated Measures ANOVA

Source	df	Λ	p	eta^2
	Between subjects			
Age Group (Age)	3,79	.852	.005	.148
Gender (G)	3,79	.965	.423	.035
Activity (Act)	6,158	.827	.019	.091
G x Act	6,158	.966	.842	.017
G x Age	3,79	.986	.766	.014
Act x Age	6,158	.915	.319	.043
G x Act x Age	6,158	.904	.230	.049
	Within subjects			
Distance (D)	12,70	.399	< .001	.631
D x G	12,70	.865	.539	.135
D x Act	24,140	.643	.098	.198
D x Age	12,70	.871	.585	.129
D x G x Act	24,140	.824	.933	.092
D x Gr x Age	12,70	.904	.816	.096
D x Act x Age	24,140	.804	.872	.103
D x G x Act x Age	24,140	.842	.966	.083

Table 5.10.

Correlations Between CPOCI Subscales and Measures of Challenge, Skill, Difficulty, Distance and Success

	CPOCI Subscales		
	Skill = Challenge	Skill > Challenge	Challenge > Skill
Challenge	.274**	-.596**	.507**
Skill	-.062	.366**	-.566**
Difficulty	.181	-.524**	.564**
Distance	.134	-.554**	.515**
Success	-.206*	.530**	-.500**

*$p < .05$; ** $p < .01$

Figure Caption

Figure 5.1. Diagram of testing area.

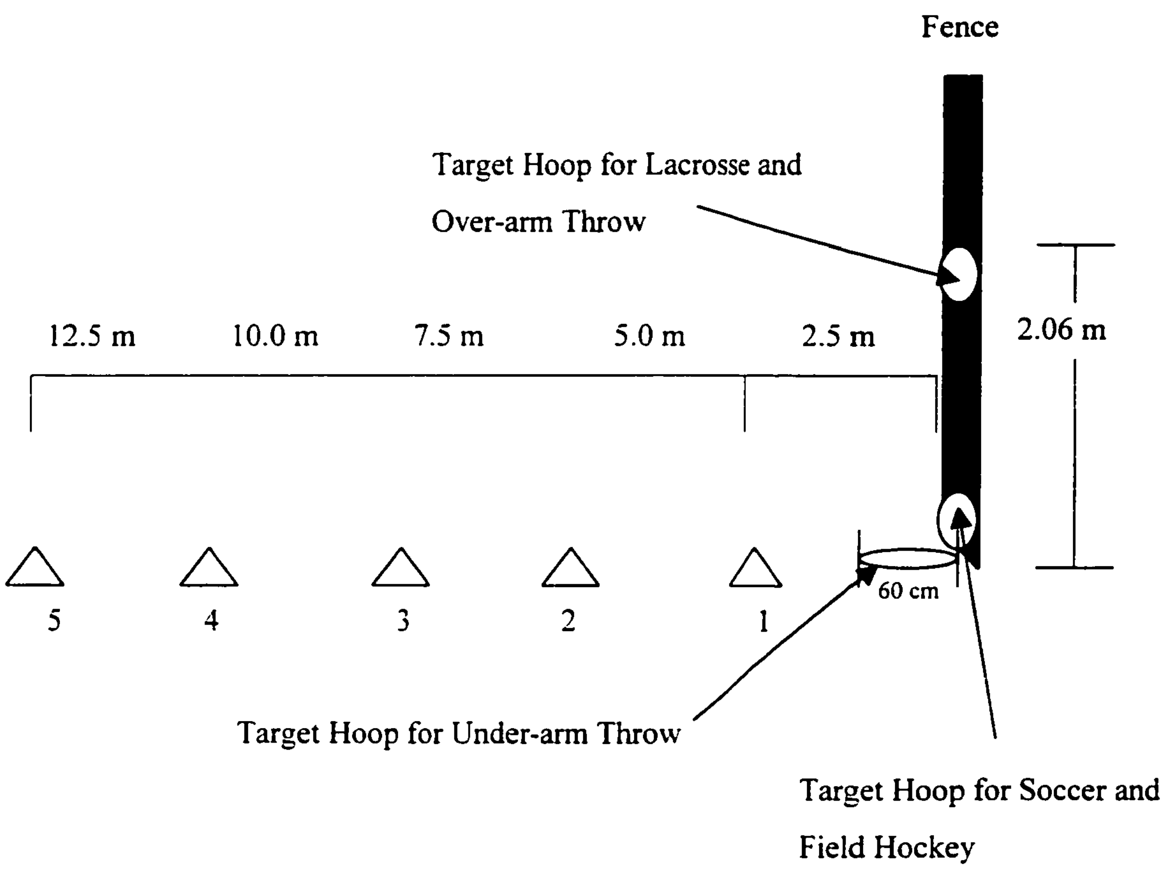

Fence
Target Hoop for Lacrosse and
Over-arm Throw
12.5 m
10.0 m
7.5 m
5.0 m
2.5 m
2.06 m
60 cm
5
4
3
2
1
Target Hoop for Under-arm Throw
Target Hoop for Soccer and
Field Hockey

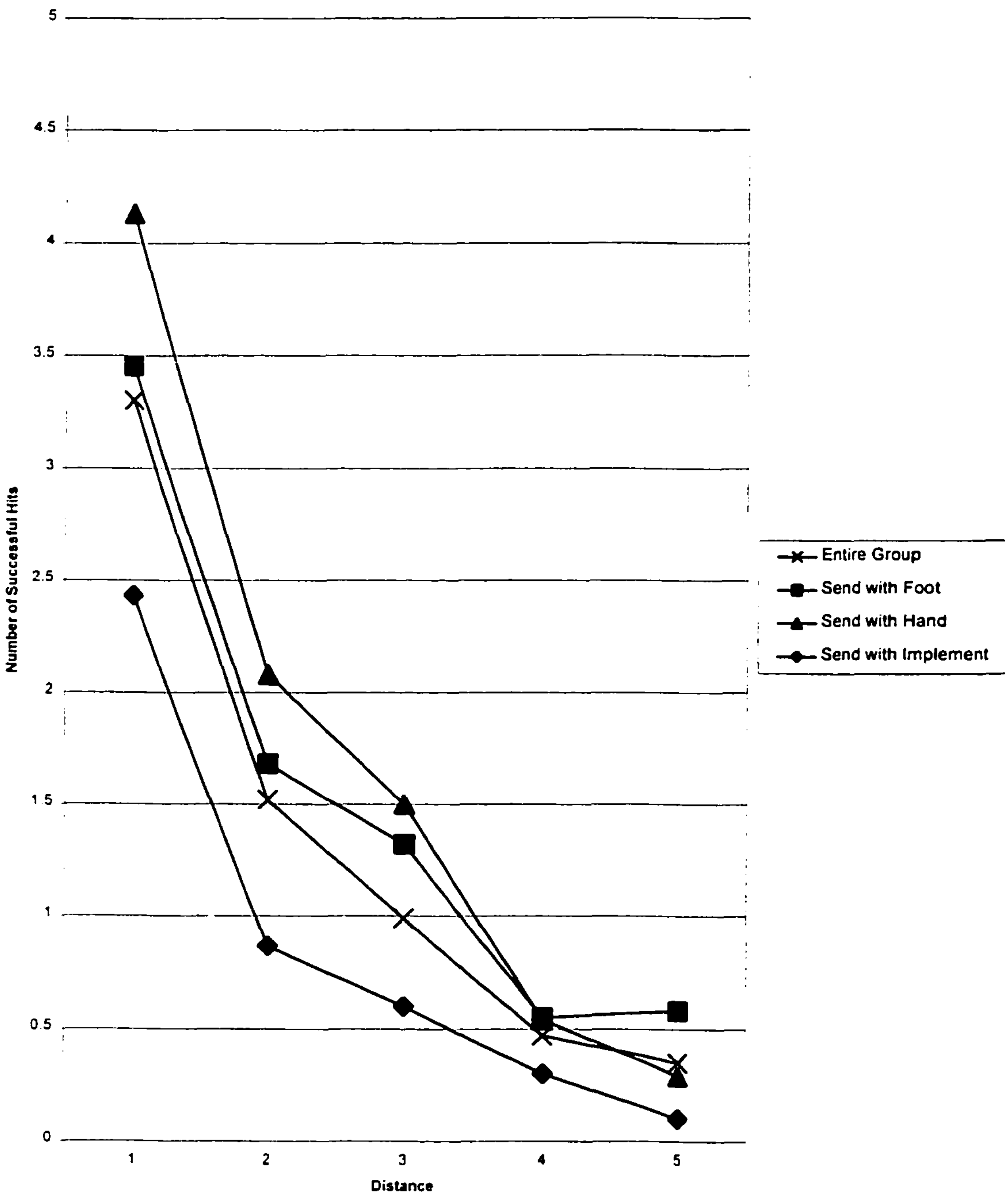
Figure Caption

Figure 5.2. Children's perceptions of optimal challenge items.

	Skill = Challenge (S = C)	Challenge > Skill (C > S)	Skill > Challenge (S> C)
1	It was in the middle of easy and hard	I needed more skill to do the activity	It was so easy, it was boring
2	It was a good challenge that I could do	It was too hard for me to do	I had more skill than I needed to do it
3	It was not too easy and not too hard	It was too challenging for my skill level	I could do it really easily
4	It was challenging, but I could still do it	I wanted to quit because it was too hard	It was simple because I had done it really good lots of times before
5	The challenge was perfect for my skill level	I was not very good at it	My skills were a lot higher than the challenge
6	My skills were equal to the challenge of the activity	The challenge was higher than what I could do	It was so easy, I did not want to do it anymore

Figure Caption

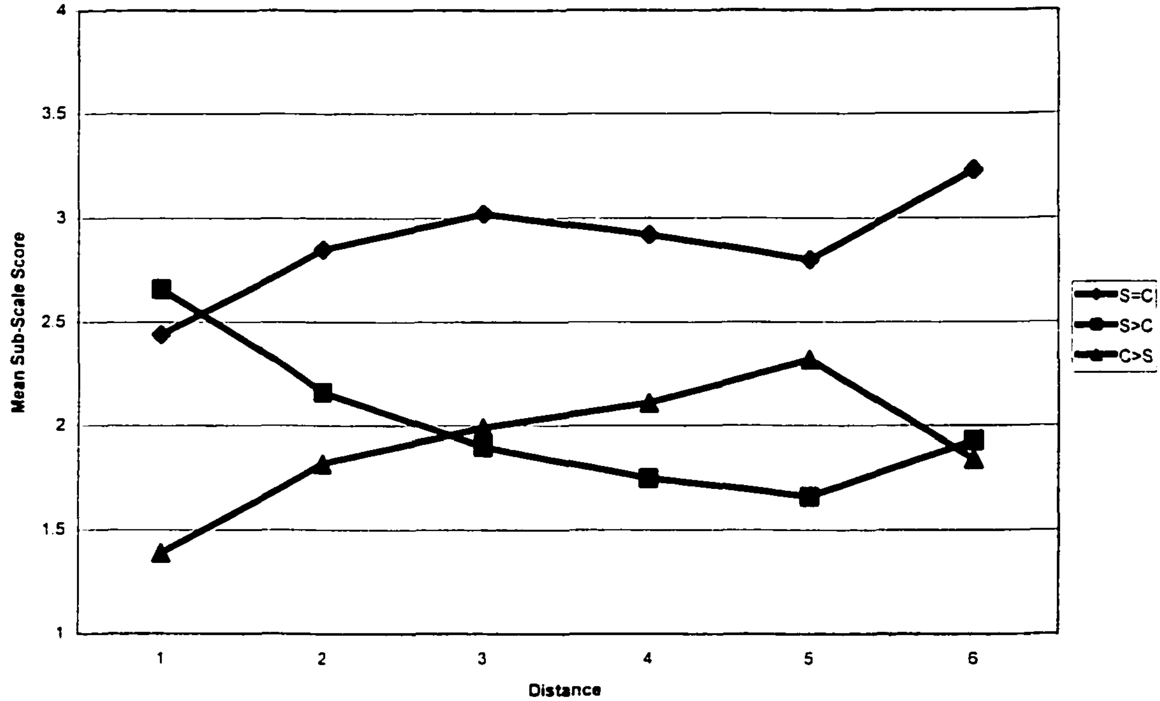
Figure 5.3. Number of successful hits per distance for entire group and by type of activity.



5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Number of Successful Hits
1
2
3
4
5
Distance
Entire Group
Send with Foot
Send with Hand
Send with Implement

Figure Caption

Figure 5.4. Group score by trial distance for entire sample.



*Note: Distance 6 = Participants chose distance that was not too easy yet not too hard

References

Arnheim, D. D., & Pestolesi, R. A. (1978). Elementary physical education: A developmental approach (2nd ed.). St. Louis, Missouri: C.V. Mosby.

Biddle, S. J. H., & Chatzisarantis, N. (1999). Motivation for a physically active lifestyle through physical education. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand, & R. Seiler (Eds.), Psychology for physical educators (pp. 5 - 26). Champaign, IL: Human Kinetics.

Boggiano, A. K., Main, D. S., & Katz, P. A. (1988). Children's preferences for challenge: The role of perceived competence and control. Journal of Personality and Social Psychology 54, 134 - 141.

Bouffard, M. (1993). The perils of averaging data in adapted physical activity research. Adapted Physical Activity Quarterly, 10, 371 - 391.

Brustad, R. J. (1998). Developmental considerations in sport and exercise psychology measurement. In J. Duda (Ed.), Advances in sport and exercise psychology measurement (pp. 461 - 470). Morgantown, WV: Fitness Information Technology.

Bryant, F. B., & Yarnold, P. R. (1997). Principal-components analysis and exploratory factor analysis. In L. G. Grimm, & P. R. Yarnold (Eds.), Reading and understanding multivariate statistics (pp. 99 - 136). Washington, DC: American Psychological Association.

Centre for Disease Control and Prevention (1996). Physical activity and health: A report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services.

Centre for Disease Control and Prevention (1997). Guidelines for school and community programs to promote lifelong physical activity among young people. Morbidity and Mortality Weekly Report, 46(RR-6), 1 - 42.

Chalip, L., Csikszentmihalyi, M., Kleiber, D., & Larson, R. (1984). Variations of experience in formal and informal sport. Research Quarterly, 55, 109 - 116.

Corbin, C. B. (1980). A textbook of motor development (2nd ed.). Dubuque, IO: Wm.C. Brown Company.

Craig, C. L., Russell, S. J., Cameron, C., & Beaulieu, A. (1999). Foundation for joint action: Reducing physical inactivity. Ottawa, ON: Canadian Fitness and Lifestyles Research Institute.

Cratty, B. J. (1967). Movement behaviour and motor learning (2nd ed.). Philadelphia, PA: Lea & Febiger.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Philadelphia, PA: Harcourt Brace Jovanovich College.

Cronbach, L. J. (1984). Essentials of psychological testing. New York: Harper & Row.

Csikszentmihalyi, M. (1975). Beyond boredom and anxiety. San Francisco, CA: Jossey-Bass.

Csikszentmihalyi, M. (1990). Flow. New York: Harper & Row.

Csikszentmihalyi, M., & Csikszentmihalyi, I. (1988). Optimal experience. Cambridge, UK: Cambridge University Press.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. The Journal of Nervous and Mental Disease, 175, 526 - 536.

Csikszentmihalyi, M., Rathunde, K., & Whalen, S. (1997). Talented teenagers. The roots of success and failure. Cambridge, UK: Cambridge University Press.

Danner, F. W., & Lonky, E. (1981). A cognitive-developmental approach to the effects of rewards on intrinsic motivation. Child Development, 52, 1043 - 1052.

Deci, E. L. (1975). Intrinsic motivation. New York: Plenum.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behaviour. New York: Plenum.

DeOreo, K., & Keogh, J. (1980). Performance of fundamental motor tasks. In C. Corbin (Ed.), A textbook of motor development (2nd ed.) (pp. 76 - 91). Dubuque, IO: Wm.C. Brown Company.

DeVillis, R. F. (1991). Scale development: Theory and application. Newbury Park: Sage.

Dobos, J. A. (1996). Collaborative learning: Effects of student expectations and communication apprehension on student motivation. Communication Education, 45, 118 - 134.

Ellis, G. D. (1987). Development of a typology of common adolescent free time activities: A validation and extension of Kleiber, Larson and Csikszentmihalyi. Journal of Leisure Research, 19, 284 - 292.

Ellis, G., Voelkl, J. E., & Morris, C. (1994). Measurement and analysis issues with explanations of variance in daily experiences using the flow model. Journal of Leisure Research, 26, 337 - 356.

Feltz, D. L., Gould, D., & Horn, T.S. (1982). Perceived competence among youth sport participants and dropouts. Manuscript submitted for publication. Michigan State University, East Lansing, MI.

Golden, C. J., Sawicki, R. F., & Frazen, M. D. (1984). Test construction. In G. Goldstein, & M. Hersen (Eds.). Handbook of psychological assessment (pp. 19 - 37). New York: Pergamon.

Harter, S. (1974). Pleasure derived by children from cognitive challenge and mastery. Child Development, 45, 661 - 669

Harter, S. (1978a). Effectance motivation reconsidered. Human Development, 21, 34 - 64.

Harter, S. (1978b). Pleasure derived from challenge and the effects of receiving grades on children's difficulty level choices. Child Development, 49, 788 - 799.

Harter, S. (1985). Manual for the self-perception profile for children. Denver, CO: University of Denver.

Harter, S. (1992). The relationship between perceived competence, affect, and motivational orientation within the classroom: Processes and patterns of change. In A. K. Boggiano, & T. S. Pittman (Eds.), Achievement and motivation: A social-developmental perspective (pp. 77 - 114). Cambridge, UK: Cambridge University Press.

Hastad, D. N., & Lacy, A. C. (1998). Measurement and evaluation in physical education and exercise science (3rd ed.). Boston, MA: Allyn & Bacon.

Linacre, J. M. (2000). New approaches to determining reliability and validity. Research Quarterly for Exercise and Sport, 71, 129 - 136.

Livingston, L. A., & Mandigo, J. L. (1999). Bilateral Q angle asymmetry and anterior knee pain syndrome. Clinical Biomechanics, 4, 7 - 13.

Mahoney, C. A., Thombs, D. L., & Howe, C. Z. (1995). The art and science of scale development in health education research. Health and Education Research, 10, 1 - 10.

Mandigo, J. L. (2001a). Listening to the voices of children: Taking the first steps in the development of a self-report instrument of children's perceptions of optimal challenge. Unpublished manuscript. University of Alberta, Edmonton, Alberta, Canada.

Mandigo, J. L. (2001b). The importance of content-relevance evidence to support the construct validity of the children's perceptions of optimal challenge instrument. Unpublished manuscript. University of Alberta, Edmonton, Alberta, Canada.

Mandigo, J. L. (2001c). Using children's similarity ratings as a source of construct validity evidence. Unpublished manuscript. University of Alberta, Edmonton, Alberta, Canada.

Mandigo, J. L., & Couture, R. T. (1996). An overview of the components of fun in physical education, organised sports and physical activity programs. Avante, 2, 56 - 72.

Mandigo, J. L., & Holt, N. L. (2000). Putting theory into practice. How cognitive evaluation theory can help us better understand how to motivate children in physical activity environments. Journal of Physical Education, Recreation and Dance, 71, 44 - 49.

Mandigo, J. L., & Thompson, L. P. (1998). Go with their flow: How flow theory can help practitioners to intrinsically motivate children to be physically active. The Physical Educator, 55, 145 - 159.

Mandigo, J. L., Thompson, L. P.,& Couture, R. T. (1998). Equating flow theory with the quality of children's physical activity experiences (Abstract). Journal of Sport and Exercise Psychology, 20(Suppl.), S94.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed.) (pp. 13 - 103). New York: American Council on Education.

Reeve, J. (1996). Motivating others: Nurturing inner motivational resources. Needham Heights, MA: Allyn & Bacon.

Roberts, G. C., Kleiber, D.A., & Duda, J. L. (1981). An analysis of motivation in children's sport: The role of perceived competence in participation. Journal of Sport Psychology, 3, 206 - 216.

Rogers, C. S., & Ponish, K. K. (1987, April). Control of level of challenge: Effects on intrinsic motivation to play. Paper presented at the Society for Research in Child Development, Baltimore, MA.

Rowley, D. A. (1996). Physical activity through the eyes of six ten-year-old children: A case study. Unpublished dissertation. University of Alberta, Edmonton, Alberta, Canada.

Rust, J., & Golombok, S. (1989). Modern psychometrics: The science of psychological assessment. London, UK: Routledge.

Ryan, R. M., Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemporary Educational Psychology, 25, 54 - 67.

Sanders, S., & Graham, G. (1995). Kindergarten children's initial experiences in physical education: The relentless persistence for play classes with the zone of acceptable responses. Journal of Teaching in Physical Education, 14, 372 - 383.

Sarrazin, P., & Famose, J. (1999). Children's goals and motivation in physical education. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand, & R. Seiler (Eds.), Psychology for physical educators (pp. 27 - 50). Champaign, IL: Human Kinetics.

Scanlan, T.K., & Simons, J.P. (1992). The construct of sport enjoyment. In G. C. Roberts (Ed.), Motivation in sport and exercise (pp. 199 - 215). Champaign, IL: Human Kinetics.

Stone, W. L., & Lemanek, K.L. (1990). Developmental issues in children's self-reports. In A. M. La Greca (Ed.), Through the eyes of the child (pp. 18 - 56). Needham Heights: Allyn & Bacon.

Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate statistics (4th ed.). Boston, MA: Allyn & Bacon.

Turner, J. C., Parkes, J., Cox, K. E., & Meyer, D. K. (1995). The role of optimal challenge in students literacy engagement. In K. A. Hinchman, D. J. Leu, & C. K. Kinzer (Eds.), Perspectives on literacy research and practice. Forty-fourth yearbook of the National Reading Conference (pp. 126 - 136).

Vallerand, R. J., & Fortier, M. S. (1998). Measures of intrinsic and extrinsic motivation in sport and physical activity: A review and critique. In J. Duda (Ed.), Advances in sport and exercise psychology measurement (pp. 81 - 101). Morgantown, WV: Fitness Information Technology.

Voelkl, J. E., & Ellis, G. D. (1998). Measuring flow experiences in daily life: An examination of the items used to measure challenge and skill. Journal of Leisure Research, 30, 380 - 389.

Wankel, L. M. (1993). The importance of enjoyment to exercise adherence and psychological benefits from physical activity. International Journal of Sport Psychology, 24, 151 - 169.

Weiss, M. R. (1986). Self-esteem and achievement in children's sport and physical activity. In D. Gould, & M. Weiss (Eds.), Advances in pediatric sport sciences: Vol. 2, (pp. 87 - 119). Champaign, IL: Human Kinetics.

Weiss, M. R., & Bressan, E. S. (1985). Relating instructional theory to children's psychosocial development. Journal of Physical Education, Recreation and Dance, November/December, 34 - 36.

Weiss, M. R., Bredemeier, B. J., & Shewchuk, R. M. (1985). An intrinsic/extrinsic motivation scale for the youth sport setting: A confirmation factor analysis. Journal of Sport Psychology, 7, 75 - 91.

CHAPTER 6

Conclusion

The purpose of the studies outlined in this dissertation was to create developmentally appropriate items for a self-report instrument of optimal challenge for use with children and to present various sources of construct validity evidence that increase the likelihood that the inferences made from such an instrument reflect children's perceptions of being optimally challenged. After going through various studies to first of all create the items by using children's descriptions and experiences, and then examining the validity of the items as they pertain to the construct of optimal challenge, a reliable instrument with various sources of validity evidence has been developed (see Appendix A). This has been demonstrated through various sources of construct validity evidence that began with content validity and worked its way through various sources of validity evidence (e.g., criterion-related, predictive, discriminant, content relevance, convergent, etc.).

In Study 1, 27 children reflected on previous physical activity experiences to describe events commonly linked to the construct of optimal challenge. The trustworthiness of the data that emerged was established through a number of methods including a comparison of the data to an alternate sample of five children. After ensuring as best as possible through numerous trustworthiness techniques that the data reflected children's actual experiences, items for a self-report instrument were generated. The wording of the items were based upon how children described and experienced various levels of optimal challenge (i.e., skill greater than challenge, challenge greater than skill and skill equals challenge). The relevance of these items was then tested by asking a team of nine expert judges to rate how well the items reflected the different levels of optimal challenge. These judges were also free to write in comments about each item. These ratings and comments were then used to modify the initial bank of items to come up with six items for three different subscales reflecting different levels of the optimal challenge construct. In order to ensure that the revised items did truly reflect children's perceptions of different levels of optimal challenge, they were reviewed by a group of children in Grades 2 and 3. Using the Conditional Rank Ordering technique, children rated the similarity between the different items. Children also sorted the items into similar piles and named them accordingly. The similarities and differences between the ratings were analysed using multidimensional scaling techniques. These analyses are similar to a factor analysis except that they allow for a direct assessment of how children perceive the similarities and differences between items. After ensuring that the items were an accurate reflection of how children would describe different levels of optimal challenge through the Sorting Technique, the items were placed into a questionnaire format that was consistent with other self-report instruments used with children. The readability of this instrument was deemed to be appropriate for children at the Grade 3 level. The psychometric properties of the instrument were then tested in a field setting that required participants to take part in an aiming task that had a number of different levels of difficulty. The results of the field test demonstrated that the instrument was reliable and various sources of construct validity were presented at both the idiographic and nomothetic levels. As a result of using this unitary approach to construct validity as recommended by Messick (1989) and using various pieces of evidence gathered through this process, an increased confidence has been developed that the observed scores generated from the Children's Perception of Optimal Challenge Instrument (CPOCI) are close to each individual's true score.

Limitations

Despite this confidence, there were some limitations to the studies conducted in this dissertation. In Study 1, participants in the replication study did not voluntarily use the word "challenge" to describe their experiences. Although the children in the replication study did understand the word challenge when asked and had similar responses to those given by the participants in the initial study, they did not use it to describe their experiences like the children in the initial study. This may have been a function of the sample size and theoretical saturation not being reached and/or a function of the researcher also being the participants' instructor with the initial group of children and influencing the type of language the children were using to describe their experiences. Future research may wish to closely examine how close researchers can be to the participants before it causes too much bias in the participants' responses. Currently, there is a wide spectrum of opinions concerning how close researchers should be when conducting qualitative research (e.g., Creswell, 1994; Jackson, 1995). Future research should also consider looking at whether children do use the word challenge to describe their experiences and a closer understanding of "how" this word is used in children's language.

In Study 2, the relevance of the items as they pertained to the three subscales was examined by the judges. Messick (1989), however, recommend that the representation of the items as a whole should also be examined to ensure that the items that make up the instrument do in fact represent all the different domains of the construct. This was not done in Study 2. Future research should go back to judges after all of their comments and ratings have been collected. After making appropriate revisions, the judges can then rate how well they feel the remaining bank represent the entire construct.

In Study 3, the main finding suggested that the items from the three different subscales were interpreted by children in the manner for which they were created. Children tended to rate items from the same subscale as being similar and to rate items from different subscales as being dissimilar. The children were also able to group items from the same subscale together. When analysing the MDS results, a two dimensional solution based on theoretical interpretations was recommended despite the three dimensional solution having better goodness-of-fit indices. Future research may wish to examine the potential of a three dimensional solution to explain how the items were grouped together. Due to the paucity of optimal challenge research, however, this is not possible at this time. As a result, theoretical interpretations of the three dimensional solution are somewhat limited. Should this body of literature progress and closer theoretical ties are made to optimal challenge, it may be possible to come up with a three dimensional interpretation of the items within the CPOCI.

In Study 4, although the subscale means did fit the hypothesised structure based on their relationship to distance from the target, there were some differences in this relationship when examining the individual profiles. The reasons for these differences, however, were not investigated. Although there may be cause for concern given the lack of inter-rater reliability (i.e., only the lead researcher analysed the individual profiles), the individual profiles did suggest some differences at the idiographic level. It is not known whether these differences were a result of different interpretations of the items between individuals or whether there were individual dispositional or situational factors influencing the responses. As well, although the factor analyses showed similar simple structures across all five trials, there were some items that did cross load thus suggesting differences in interpretation. Future research may wish to go back to participants and ask them for their interpretations. These interpretations of the items can then be used to help explain the individual profiles to determine whether there are some items that have different interpretations among children. This may also help address some concerns regarding potential differences between the type of language children use to describe a construct and the type of language adults use to describe the same construct.

One other limitation to the studies in this dissertation is the lack of an older age group in the development phase of the items. All of the participants in Studies 1 and 3 were from Grades 2 or 3. However, in Study 4, participants as old as 12 years of age (Grade 6) were included in the study. Future research may wish to replicate the first three studies with children in the older age group (i.e., ages 10 – 12 years) to compare potential differences in the way this age group describes, experiences, and prefers different levels of optimal challenge.

Future Research and Practical Uses for the CPOCI

Validity is a never-ending process (Messick, 1989). Therefore, future research programs that utilise this instrument should continue to provide further evidence (e.g., confirmatory factor analysis) and take into consideration recommendations presented in the final study as outlined in Chapter 5 (e.g., age group differences, perceived competence, and experience differences) and the limitations previously described. Should the psychometric properties presented from these and subsequent studies remain consistent, then this only serves to strengthen the evidence and interpretations made as a result of using the CPOCI. Future research and educational uses of the CPOCI can then commence to help further the knowledge surrounding the optimal challenge construct.

Potential uses of CPOCI for educational purposes. The CPOCI may have potential benefits for classroom teachers. Teachers are constantly trying to create developmentally appropriate activities. These activities are characterised by taking individual and group abilities into consideration and structuring activities that match these abilities (Bredekamp, 1992). Based on this description, it could be argued that developmentally appropriate activities are optimally challenging. Therefore, teachers could use the CPOCI to rate the degree to which the activities and learning environment they create are optimally challenging. If students continually score low on the optimal challenge subscale and high on one of the other two subscales (i.e., skill higher than challenge or challenge higher than skill), they may want to consider increasing or decreasing the challenge of the activity so that it matches the skill level of their students. Teachers can then take this information to help continually expose children to optimally challenging activities by looking at individual profiles. There may be some in the class who are being optimally challenged, but others who are either continually bored or frustrated within a physical education setting as a result of a skill and challenge imbalance. The CPOCI may help identify students' degree of optimal challenge and allow teachers to give more specific and individualised feedback and tasks to such students.

Potential use of CPOCI for research purposes. The CPOCI has great potential in helping to bridge the gaps between different motivational theories. For example, all three humanistic theories identified previously (Csikszentmihalyi, 1990; Deci & Ryan, 1985; Harter, 1978) allude to the importance of optimal challenge and enjoyment. Fun has also been defined by some as the balance between perceived skill and challenge (e.g., Briggs, 1994; Petlichkoff, 1992). Participants in Studies 1, 3, and 4 did indicate that when they felt their skills and the challenge of the activity were equal, that it was fun/ enjoyable. Mandigo and Couture (1996) have also found evidence to support these claims. They found that when children reported high skill balanced with high challenge, children reported significantly high levels of fun. However, concerns have since been raised over the use of measures of fun, skill, and challenge as used by Mandigo and Couture (1996) and many others (e.g., Harter, 1974; Stein, Kimiecik, Daniels & Jackson, 1995) that lack psychometric evidence. Now that a measure of optimal challenge has been developed, future research can use the CPOCI and determine the relationship between levels of optimal challenge and fun as measured by existing motivational instruments developed and validated for children in physical activity environments (e.g., Intrinsic Motivation Inventory [Whitehead & Corbin, 1991]). This may help to legitimise the importance of fun in physical education by linking it to the construct of optimal challenge given that some have seen "fun" as something frivolous and not linked to educational outcomes (Whitehead, 1988).

The CPOCI can also be used to examine the relationship that optimal challenge has to other motivational constructs. Based upon Deci and Ryan's (1985) self-determination theory and Harter's (1978) competence motivation theory, when an individual is continually optimally challenged, it can lead to positive psychological outcomes such as increased levels of self-determination, autonomy, competence, and self-efficacy. In addition, the CPOCI can be used to help determine potential situational antecedents that facilitate optimal challenge. Theoretical models and empirical research suggest that situational variables such as teaching styles (Mosston, 1992), and the functional significance of rewards and marks (Harter, 1974; Ryan & Deci, 2000) can have an impact on creating optimally challenging experiences. For example, an inclusive teaching style as indicated by Mosston (1992) can help facilitate a balance between skill and challenge because children choose the level they want to start working on for a task. As well, dispositional characteristics such as a child's goal orientation (Sarrazin & Famose, 1999), perceptions of competence (Harter, 1978), and motivational orientation (Harter, 1981) have been suggested to influence the degree to which a child will be optimally challenged. The CPOCI can be used in experimental studies to determine the impact that these theoretically linked antecedents have on different levels of optimal challenge.

Closing Remarks

In an era where concerns across the globe have been raised about ways to encourage children to be more physically active, there is an increased need to find ways that will attract and keep children motivated. The development of the CPOCI is simply the start of a research program. Now that a psychometrically sound instrument has been developed and the initial evidence has demonstrated that interpretations resulting from it are valid, researchers examining situational (e.g., teaching styles, functional significance of rewards) and dispositional (e.g., perceptions of competence, goal orientations) influences on optimal challenge can be conducted. The results from these studies will be beneficial in assisting all those who work with children in physical activity environments so that the chance of structuring optimally challenging activities is increased. As a result, research can also examine the long term impact that optimal challenge has on children's motivation in physical activity. For example, how often do children have to be continually optimally challenged before it has a long term impact on their desire to be active. Conversely, how many episodes of not being optimally challenged (i.e., skill ≠ challenge) does it take before children are turned off of physical activity for a lifetime. These type of long term effects of optimal challenge on intrinsic motivation are not known despite the potential it may have for increasing the chances of creating an active generation.

Notwithstanding some of the limitations to the findings from the series of studies outlined in this dissertation, researchers now have an instrument at their disposal to use and improve upon in order to tap into children's levels of optimal challenge. Although this will take time to collect further validation evidence and to examine the antecedents and short and long term impact of optimal challenge, the research has to start somewhere. It is hoped that by going directly to children and finding out how to structure activity environments that optimally challenge them, this will be the last generation of inactive children. Continually exposing children to programs that are aimed at providing optimally challenging activities, will only serve to enhance enjoyment and motivation for children to be active for a lifetime.

References

Bredekamp, S. (1992). What is "developmentally appropriate" and why is it important? Journal of Physical Education, Recreation and Dance, 63(6), 31 - 32.

Briggs, J. D. (1994). An investigation of participant enjoyment in the physical activity instructional setting. The Physical Educator, 51, 213 - 221

Creswell, J. W. (1994). Research design. Qualitative and quantitative approaches. London, UK: Sage.

Csikszentmihalyi, M. (1990). Flow. New York: Harper & Row.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. New York: Plenum

Harter, S. (1974). Pleasure derived by children from cognitive challenge and mastery. Child Development, 45, 661 - 669.

Harter, S. (1978). Effectance motivation reconsidered. Human Development, 21, 34 - 64.

Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. Developmental Psychology, 17, 300 - 312.

Jackson, S. A. (1995). Factors influencing the occurrence of flow states in elite athletes. Journal of Applied Sport Psychology, 7, 138 - 166.

Mandigo, J. L., & Couture, R. T. (1996). An overview of the components of fun in physical education, organised sports and physical activity programs. AVANTE, 2, 56 - 72.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed.) (pp. 13 - 103). New York: American Council on Education.

Mosston, M. (1992). Tug-O-War, no more: Meeting teaching-learning objectives using the spectrum of teaching styles. Journal of Physical Education, Recreation, and Dance, 62(1), 27 - 31; 56

Petlichkoff, L. M. (1992). Youth participation and withdrawal. Is it simply a matter of fun? Pediatric Exercise Science, 4, 105 - 110.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55, 68-78.

Sarrazin, P., & Famose, J. (1999). Children's goals and motivation in physical education. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand, & R. Seiler (Eds.), Psychology for physical educators (pp. 27 - 50). Champaign, IL: Human Kinetics

Stein, G. L., Kimiecik, J. C., Daniels, J., & Jackson, S. A. (1995). Psychological antecedents of flow in recreational sport. Personality and Social Psychology Bulletin, 21, 125 - 135.

Whitehead, J. R., & Corbin, C. B. (1991). Effects of fitness test type, teacher, and gender on exercise intrinsic motivation and physical self-worth. Journal of School Health, 61, 11 - 16.

Whitehead, M. (1988). The danger of fun. British Journal of Physical Education, 19(4), 155.

Appendix A

Children's Perceptions of Optimal Challenge Instrument (CPOCI).

1. Only think about the activity you just did.
2. Read the question in the middle box.
3. Decide whether it is True or False for you and then stamp the box.
4. Then stamp the box beside it that is the right answer for you.
5. Try to answer each question as quickly as possible and give the answer that is right for you.

For example:

		False	ABOUT THE ACTIVITY I JUST DID ...	True		
Really False for Me	Sort of False for Me	☐	It was fun to do	☐	Sort of True for Me	Really True for Me

Remember this is not a test. There are no wrong answers. Put down the right answer for you!

****

			False	ABOUT THE ACTIVITY I JUST DID ...	True		
1.	Really False for Me	Sort of False for Me	☐	It was in the middle of easy and hard	☐	Sort of True for Me	Really True for Me
2.	Really False for Me	Sort of False for Me	☐	I needed more skill to do the activity	☐	Sort of True for Me	Really True for Me
3.	Really False for Me	Sort of False for Me	☐	It was so easy, it was boring	☐	Sort of True for Me	Really True for Me
4.	Really False for Me	Sort of False for Me	☐	It was a good challenge that I could do	☐	Sort of True for Me	Really True for Me
5.	Really False for Me	Sort of False for Me	☐	It was too hard for me to do	☐	Sort of True for Me	Really True for Me

			False	ABOUT THE ACTIVITY I JUST DID …	True		
6.	Really False for Me	Sort of False for Me	☐	I had more skill than I needed to do it	☐	Sort of True for Me	Really True for Me
7.	Really False for Me	Sort of False for Me	☐	It was not too easy and not too hard	☐	Sort of True for Me	Really True for Me
8.	Really False for Me	Sort of False for Me	☐	It was too challenging for my skill level	☐	Sort of True for Me	Really True for Me
9.	Really False for Me	Sort of False for Me	☐	I could do it really easily	☐	Sort of True for Me	Really True for Me
10.	Really False for Me	Sort of False for Me	☐	It was challenging, but I could still do it	☐	Sort of True for Me	Really True for Me
11.	Really False for Me	Sort of False for Me	☐	I wanted to quit because it was too hard	☐	Sort of True for Me	Really True for Me
12.	Really False for Me	Sort of False for Me	☐	It was simple because I had done it really good lots of times before	☐	Sort of True for Me	Really True for Me
13.	Really False for Me	Sort of False for Me	☐	The challenge was perfect for my skill level	☐	Sort of True for Me	Really True for Me
14.	Really False for Me	Sort of False for Me	☐	I was not very good at it	☐	Sort of True for Me	Really True for Me
15.	Really False for Me	Sort of False for Me	☐	My skills were a lot higher than the challenge	☐	Sort of True for Me	Really True for Me

			False	ABOUT THE ACTIVITY I JUST DID ...	True		
16.	Really False for Me	Sort of False for Me	☐	My skills were equal to the challenge of the activity	☐	Sort of True for Me	Really True for Me
17.	Really False for Me	Sort of False for Me	☐	The challenge was higher than what I could do	☐	Sort of True for Me	Really True for Me
18.	Really False for Me	Sort of False for Me	☐	It was so easy, I did not want to do it anymore	☐	Sort of True for Me	Really True for Me