

**ESTABLISHING A VALID AND RELIABLE COMPETENCY ASSESSMENT
FOR GYMNASTICS COACH CERTIFICATION**

A Thesis Submitted to the College of Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the

Degree of Masters of Science

In the

College of Kinesiology

University of Saskatchewan

Saskatoon, Saskatchewan

By

Alynn C. Brischuk

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other uses of materials in this thesis in whole or part should be addressed to:

Head of the Department of Kinesiology
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5B2
Canada

OR

Dean
College of Graduate Studies and Research
University of Saskatchewan
107 Administration Place
Saskatoon, Saskatchewan S7N 5A2
Canada

ABSTRACT

Coaches in Canada are trained through the National Coaching Certification Program (NCCP) administered by The Coaching Association of Canada. The NCCP is the national standard and recently shifted its educational emphasis from the transfer of knowledge to the development of coaching competency. As a result, coaches are required to demonstrate competency in a specified sport and level prior to being awarded a coaching certification. The purpose of this study was to establish validity and reliability of an NCCP coaching competency assessment.

Gymnastics Canada agreed to have their Community Sport coaching competency instrument tested for content validity, face validity, inter-rater reliability and intra-rater reliability. Their original assessment instrument was revised during content validity testing based on recommendations from five coaching context experts. Removing six items, adding one item, and revising four items for clarity modified the original 48-item instrument. An assessment of overall competency was added to the instrument to rate the coach as Beginner, Competent, Proficient, or Expert. Three certified coaches confirmed the face validity of the modified instrument.

Reliability tests were conducted on the ratings provided by ten experts who observed a coach's lesson on video. The result was a moderate level of inter-rater reliability, displayed by an Agreement Coefficient (AC_1) of 0.43 and a Percent of Agreement (PA) of 72%. Nine of the ten raters assessed the coach's performance as Competent, Proficient or Expert, while one of the ten rated the coach as Beginner. After a repeat observation of the same coaching performance, the intra-rater reliability of five raters resulted in agreement levels of *Moderate* ($AC_1 = 0.45$; $PA = 67\%$), *Substantial* ($AC_1 = 0.75, 0.77$; $PA = 81\%, 89\%$), and *Almost Perfect* ($AC_1 = 0.87, 0.82$; $PA = 93\%, 90\%$).

The modified instrument has content and face validity. However, its usefulness is impacted by variability in rater preferences. The inter-rater reliability results attest to concerns about the consistency in assessment of a coaching certificate for Community Sport Artistic Gymnastics coaches. Suggestions to improve the reliability of the instrument include training raters to classify the coach's competency on a standardized scale of sport-specific expertise.

ACKNOWLEDGEMENTS

I would like to thank Dr. Keith Russell for his inspiration, mentorship and thoughtful care for my educational and professional development.

Also to the Coaching Association of Canada for providing funding of this project as a portion of a multi-sport review of coaching competency assessments and for their support of coaching research in Canada.

A sincere thank you to Gymnastics Canada, particularly Ms. Elisabeth Bureau and Ms. Suzanne Fisher for their assistance in approving this study and recruiting experts in gymnastics coach education. To all of the Learning Facilitators who gave their time and expertise, it is because of you that this study was possible and I could not be more appreciative.

To Gymnastics Saskatchewan and the regional coaches who not only supported me through my educational leave to complete this milestone in my life, but also provided access to their coaching environments and personal experiences.

Finally, to my many gymnastics friends and colleagues who continue to show me how amazing our sport is and that perfection is worth striving for.

DEDICATION

To my family, you are my strength, and my support and you never waiver in your belief of me - I love you all. And to Robert, without you I would not have had the courage to venture down this path. Your encouragement to focus on living in the moment and your convincing trust that I can do anything are my sources of hope.

TABLE OF CONTENTS

	Page
PERMISSION TO USE	i
ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
DEDICATION	iv
TABLE OF CONTENTS	v
LIST OF TABLES	vii
LIST OF FIGURES	viii
GLOSSARY	ix
1. INTRODUCTION	1
2. REVIEW OF LITERATURE	
2.1. The Role of Coaches in Sport	2
2.2. Coach Education	
i. Learning To Be a Coach	2
ii. Effective Coaching	4
iii. Expert Coaches	5
iv. Is Coaching a Profession?	6
v. The Vocation of Coaching	7
vi. International Trend of Sport Coach Development	8
2.3. Competency-Based Training	
i. Competency in Human Resource Management	9
ii. Competency Assessment	11
iii. Competency of Sports Coaches	12
2.4. Testing Reliability and Validity of CBET Programs	14
i. Reliability	15
ii. Validity	16
2.5. Purpose of the Study.....	18
3. METHODS & RESULTS	
3.1. Instrument	19
3.2. Five Stage Process	19
3.3. Stage 1 Rater Orientation	
i. Procedures	20
ii. Stage 1 Results	23
3.4. Stage 2 Content Validity Testing	
i. Procedures	24
ii. Stage 2 Results	25
3.5. Stage 3 Face Validity Testing	
i. Procedures	29
ii. Stage 3 Results	30
3.6. Stage 4 Inter-Rater Reliability	
i. Procedures	30
ii. Stage 4 Results	31
3.7. Stage 5 Intra-Rater Reliability	
i. Procedures	32

ii. Stage 5 Results	33
4. DISCUSSION	
4.1. Stage 1	35
4.2. Stage 2	
i. Validity	36
ii. Content Validity	36
iii. Assessment Scales	37
iv. Competency Assessment	38
v. Expert Coach Participants	39
4.3. Stage 3	39
4.4. Stage 4	40
4.5. Stage 5	42
4.6. Future Directions of Coaching Competency	
i. Coaching Development	43
ii. Competency Assessment	44
4.7. Limitations	45
4.8. Future Research	46
5. CONCLUSION	48
6. REFERENCES	51
7. APPENDICES	
Appendix A: Learning facilitator letter & consent form.....	61
Appendix B: Coach letter & consent form	65
Appendix C: Athlete consent form	69
Appendix D: Stage 1 instructions and data collection form	71
Appendix E: Stage 2 instructions and content validity index	79
Appendix F: Modified instrument for Community Sport artistic gymnastics	85
Appendix G: Stage 3 instructions and face validity survey	90
Appendix H: Stage 4 instructions	94
Appendix I: Stage 4 & 5 data collection form	102

LIST OF TABLES

	Page
Table 1. Strength of Agreement on a Benchmark Scale.	21
Table 2. Stage 1 Results. The Levels of Agreement Between Five Raters and a Master Learning Facilitator.	32
Table 3. Stage 2 Results. The Number of Expert Ratings by Item.	34
Table 4. Stage 2 Responses to Content Validity Survey Question (b) by Five Raters.	37
Table 5. Stage 3 Qualitative Responses to a Face Validity Survey of the Modified Instrument.	40
Table 6. Linear Weights of Inter-Rater Reliability Coefficients.	41
Table 7. Stage 4 Results. Inter-Rater Reliability of Ten Raters.	42
Table 8. Results of Coaching Expertise Assessment by Ten Raters.	43
Table 9. Stage 5 Results. Intra-Rater Reliability for Five Raters.	44
Table 10. Coaching Expertise Assessment by Five Raters Over Repeat Observations.	44

LIST OF FIGURES

	Page
Figure 1. Performance is dependent on an individual's underlying knowledge, skills and personal characteristics.	14
Figure 2. A five-stage process for establishing validity and reliability of a gymnastics coach observation instrument (Brewer & Jones, 2002).	27

GLOSSARY

Competency – An individual’s ability to perform a role or act in an appropriate way, such as for employment. It is based on an interrelationship of the knowledge, skills, attitudes and personal attributes of the individual.

Assessment – A process of collecting evidence to determine if competency exists. Evidence may be collected with single or multiple instruments depending on the attribute being assessed.

Instrument – A test device or document used during the assessment.

Item – A required task or competency on the assessment instrument.

Rating – A quantitative value assigned to an observed attribute.

Validity – The degree to which an instrument measures what it is designed to measure.

Reliability – The degree to which an instrument is trustworthy to produce the same results over multiple assessments.

CBET – Competency-Based Education and Training

CAC – Coaching Association of Canada

NCCP – National Coaching Certification Program

ICCE – International Council for Coach Education

1. INTRODUCTION

Coaches have an important responsibility to develop athletes and lead them to success in their sport. It is common for sports organizations to require coaches to complete a training program in sport-specific knowledge and skills. In Canada, coaches access formal coach education through the National Coaching Certification Program (NCCP), and receive certification specific to the sport and level following a successful assessment of their coaching performance.

Since the inception of the NCCP, coach education has primarily been conducted through knowledge-based courses, but due to a growing interest in sports coaching as a profession, the concept of coaching competency has emerged as the primary qualification of candidates for certification. Competency is best described as the collection of knowledge, skills and personal characteristics required for an individual to have success in a position (Lucia & Lepsinger, 1999). It is assessed through a collection of evidence about the individual, and is typically based on an observation of their performance. Although competency is commonly assessed in other professional sectors such as Human Resources, Medicine, and Education (Blank, 1982; McClelland, 1973), it is a relatively new concept in training and certification of coaches.

There is a lack of academic research on the assessment of coaching competency, and the shift from knowledge-based to competency-based coach education has an array of challenges. The recent increase in use of competency assessments raises the question of how to accurately and consistently evaluate the competency of coaches. Given the complexity of coaching, it is unreasonable to require coaches to be assessed on a single set of competencies applicable to all sports at all levels. Therefore, sport specific competency assessments designed by each of the Canadian national sport organizations are required to focus on individual sport demands. The primary aim of this research is to address the absence of academic and scientific support for competency assessments of sports coaches. To achieve this, a previously published five-stage process was used (Brewer & Jones, 2002) to determine the validity and reliability of the Gymnastics Canada Community Sport competency assessment instrument.

2. REVIEW OF LITERATURE

2.1 The Role of Coaches in Sport

One can participate in sport as an athlete, a coach, an official, or an administrator, but it is the athletes who are the primary focus in any sport and attract the attention of the public, the media, and by extension, the academic community. Behind every athletic performance, however, is an individual coach or group of coaches who have instilled a base of knowledge and strategy (Becker, 2009). The influences of the coach are essential to the athlete at all stages of development (Côté & Gilbert, 2009), but are often underrated.

Sport coaching is a complex process that is a challenge to define (Jones, 2007). Part of the difficulty is in understanding the roles coaches fill, ranging from casual volunteer to full time professional (Duffy et al., 2011). Turner, Nelson, and Potrac (2012), as well as Cushion et al. (2010), describe coaching as a dynamic, socio-cultural process that requires technical instruction and transfer of knowledge in individual sport environments. Coaches require technical and tactical sport-specific knowledge, and general instructional methods, all of which are crucial components of the “coaching process” (Abraham, Collins, & Martindale, 2006). Effective coaches are able to consistently apply their knowledge and improve athlete outcomes while balancing their participants’ maturational and technical development (Côté & Gilbert, 2009).

Coaching is integral to athlete success, and this fundamental relationship applies to both competitive and recreational contexts (Côté & Gilbert, 2009; Lyle & Cushion, 2010; Sport Canada, 2012). Successful coaches recognize the differences of athletes in these two contexts in terms of goals, motivations, training and competition needs (Côté & Gilbert, 2009). A basic understanding of differences in athlete needs is imperative to coach education and allows coach educators to approach educational and training opportunities essential for success.

2.2 Coach Education

i. Learning To Be a Coach. Systematic coach education exists in many countries and is often delivered by a national sport organization or an educational institution such as a college or university (Duffy et al., 2011; Trudel, Gilbert, & Werthner, 2010). Formal training of this nature usually requires the organization to deliver standardized programs that have pre-assigned

learning outcomes. The formality of the program requires it to be curriculum driven so that upon completion, a recognized coaching certification can be awarded.

Learning to coach requires both acquisition of knowledge and active engagement in the practice of coaching (Abraham & Collins, 1998; Cushion et al., 2010; Erickson et al., 2008). This learning process is not static but rather a continuous flow of experiences, attempts, and evaluations (Cassidy, Jones, & Potrac, 2009). To understand how coaches accumulate knowledge and experience, researchers have compared coach education to existing learning models (Cushion et al., 2010; Erickson et al., 2008; Werthner & Trudel, 2006). Specifically, Cushion et al. (2010) compared existing theories on how coaches learn to coach. Cushion et al. defines learning as a relatively permanent change in the coach's knowledge, skills, attitudes or behaviours. They discovered the two leading ways coaches learn are through interactions with other coaches, and through their own practice.

There are several theories focused on how coaches learn to coach. Another prominent theory categorizes coach education as a result of formal, informal and non-formal sources (Nelson, Cushion, & Potrac, 2006). Formal sources include structured coach education or curriculum programs; informal sources relate to previous athlete experiences or to watching and interacting with peers; and non-formal sources are self-directed options such as clinics or workshops. From this theory, findings have determined that coaches desire, and are encouraged to use, learning sources that are relevant to their sporting context (Cassidy & Kidman, 2010; Gilbert, Côté, & Mallett, 2006). Therefore coaches, as learners, should seek the most relevant and practical means of acquiring knowledge and experience (Nelson, Cushion, & Potrac, 2013).

There remains a need for research into the optimal combination of learning sources for coaches. In one study, coaches in the recreational context (those working outside of elite or competitive contexts) indicated their most valued source of learning was through informal sources (Erickson et al., 2008). Competitive coaches indicated formal training was their most important learning source. The same study reported that coaches in all contexts used a combination of methods from all three learning sources. Therefore, like many professions, coaches require a blend of experiences and sources of learning. And, since individuals can determine their own education needs and tailor their preferences for learning (Erickson, Côté, & Fraser-Thomas, 2007), it is not surprising that a "best practice" of educating coaches has yet to

be found (Gilbert, 2006).

ii. Effective Coaching. Coaches and educators understand that having access to a variety of learning sources is beneficial. Simply having knowledge is not enough for a coach to become effective. Coaches must apply their knowledge to meet athlete needs in their level of either recreational sport or competitive sport. This impact of the coach on the athlete is defined as coaching effectiveness (Côté & Gilbert, 2009).

An effective coach is often misinterpreted with a successful coach. The conventional descriptions of successful coaches being the most experienced or with the best competition outcome is no longer supported (Côté & Gilbert, 2009; Ford, Coughlan, & Williams, 2009; Wiman, Salmoni, & Hall, 2010). Therefore, coaching effectiveness, rather than a performance record, is what athletes prefer. This group has voiced that they value their relationship with their coach, the athletic environment, and the coach's behaviours above competition results (Becker, 2009; Côté & Sedgwick, 2003). Coaches have also been found to prefer the definition of effectiveness, as it credits their personality, their leadership, and their organizational skills (Vallée & Bloom, 2005).

To become an effective coach requires practical experience in problem solving, decision-making, and reflection (Nash & Sproule, 2011). Coaches, therefore, not only require context specific knowledge, but also practice in applying their knowledge to influence athletes and their performance. It is known that coaches value formal training sources such as comprehensive national coach education programs to acquire knowledge (Erickson et al., 2008; Misener & Danylchuk, 2009). However, coach education needs to include extensive practice and experience to allow coaches to develop critical thinking and problem solving skills (Cushion, Armour, & Jones, 2003). Unfortunately, coaches spend countless hours in planning, in training, and at competitions, but only a small amount of time in formal education (Gilbert & Rangeon, 2011). With limited formal training, coaches are left to make decisions based on their own athletic experience and through experimentation (Lemyre, Trudel, & Durand-Bush, 2007).

It cannot be determined if learning through knowledge-based courses has had an impact on the coach's effectiveness until the coach has had an opportunity to act and reflect on their actions (Nelson & Cushion, 2006). And, since coaches have expressed that they prefer practical sport-specific training experiences to theory courses (Lemyre et al., 2007), it seems reasonable

that a combination of informal and formal learning sources is efficient training. This would produce desired outcomes based on learning sport specific competencies and would allow individuals the time they need to master them (Blank, 1982). The challenge for coach educators is to move from research into practice. They must meet standards imposed by funding agencies and sport organizations (Duffy et al., 2011) while accommodating individual preferences of learning, which are primarily informal sources (Gilbert, 2009).

iii. Expert Coaches. Inquiries into the general development of expert human performances across many domains found a broad commonality: that they had extensive experience spanning many years and thousands of hours (Ericsson & Lehmann, 1996). However, a variety of skill levels emerged in a cohort of expert performers despite similar time spent studying or practicing, leading us to believe that the 10,000 hour rule, advocated by Ericsson, is more a myth than a developmental inevitability (Hambrick et al., 2013). In studying expertise development of coaches, Côté & Gilbert (2009) found the coaches who consistently displayed effectiveness could be described as experts. The challenge for coach educators is to characterize the behaviours descriptive of expert coaches and to incorporate those behaviours into meaningful educational activities (Erickson, et al., 2007; Wiman et al., 2010).

Researchers in educational psychology who were interested in how to improve teaching for the purpose of improving student achievement, and how to differentiate the behaviours between novice and expert professionals, were able to identify specific differences that were more meaningful than simply effort or experience (Berliner, 1988; Blank, 2013; Schempp, Tan, Manross, & Fincher, 1998). Given that teaching and sports coaching have many similarities, Schempp, McCullick and Mason (2006) defined the developmental stages of becoming an expert coach following the work of Berliner (1988) on developing expertise in teaching. They describe experts as having consistently outstanding performance, requiring an extensive base of knowledge and experience. Experts are able to utilize the coaching environment effectively and synthesize their knowledge into meaningful activities and information for their athletes. In addition, experts are both highly knowledgeable and skilled in applying their knowledge. They make decisions based on processing the most important information relative to the situation and relating the current problem to previous experience.

DeMarco and McCullick (1997) described how the combination of effectiveness and

experience were insufficient on their own, but together could develop expertise. Effective coaches are able to achieve positive outcomes for athletes in the desired sport context. If a coach gains experience in one sport context, whether recreational or competitive, expertise can be attained. This is an important consideration, as coaches working in different contexts will require different knowledge base, skills and attitudes (Wiman et al., 2010). Unfortunately, the majority of research on coach education targets elite and competitive coaches (Cushion et al., 2010), even though the stages early of developing coaching expertise are relevant to both competition and recreation contexts.

iv. Is Coaching a Profession? For sports coaching to be a recognized profession, an understanding of the educational procedures, practices, and performance standards of other professions is required (Duffy et al., 2011). The common criteria defining a profession include a specific knowledge base, a unique skillset, and specialized attributes or values. Established professions also have: a duty to serve the public, a set of standards for accreditation and opportunities for continuing professional development (Buhai, 2012). Although coaches are viewed as key contributors to program and athlete success (Taylor & Garratt, 2010a), they have traditionally been volunteers. As a result, adaptations within sport organizations for planning, policies, and resource allocation for professional development of coaches is lacking.

Coaching research has looked at coach education, coach expertise, and coaching behaviours with little consideration about their link to the professionalization of coaches (Lyle & Cushion, 2010). For coaches to receive professional status, an agreement is needed between international sport organizations on standardized training and career development pathways. The agreement must reflect the essential knowledge, practical skills, and accepted values/attributes coaches need to be considered professionals. The International Council for Coach Education (ICCE) is facilitating cross-border discussions on the profession of coaching (ICCE, 2012), but each nation remains responsible to create and deliver their own pathway for training, and accreditation. Duffy et al. (2011) report:

“Sport coaching is weak in comparison with other established professions, with no widespread application of a ‘right to practice’, sparse legislative arrangements and fragmented career structures. This scenario is further reflected in sporadic and inconsistent approaches to the licensing and registration of coaches supported by structured processes for continuous professional development” (p.108).

Due to these conditions, coaching is a “blended professional area, which recognizes the respective roles of professional coaches, volunteers and pre-coaches, supported by a process of professionalization” (Duffy et al., 2011, p.118). Not every nation prefers to recognize coaches as professionals (Taylor & Garratt, 2010a), nor will all coaches consider themselves professionals. Even though differences in occupational roles of coaches exist, coaches and supporters agree that being an effective coach and meeting the athlete’s needs requires knowledge and skills in a specific sport and context (Côté & Gilbert, 2009).

v. The Vocation of Coaching. Formal coach education programs such as the United Kingdom Coaching Certificate, the Canadian National Coaching Certification Program and the Australian National Coach Accreditation Scheme are designed to award coaches with an accreditation at the completion of a training program (Mallett et al., 2009). A licensing requirement, such as these accreditations, is positive for coaching to be recognized as a vocation and a good step towards being recognized as a profession, since professions have standardized process for learners to attain a professional license (Buhai, 2012). The knowledge acquired by a coach through formal education is verified through the accreditation they receive, however, as Lemyre et al. (2007) explain, “... the acquisition of knowledge does not ensure competency on the field” (p.204). Coach education programs are encouraged to provide problem solving experiences to allow coaches to experiment and evaluate their decisions in the appropriate context (Cushion et al., 2003; Gilbert & Trudel, 2001).

This trend towards competency-based coach education instead of the traditional knowledge-based coach education is occurring internationally (ICCE, 2010). By embracing this new structure, sport organizations align the desire of learners to gain practical experience (Nelson et al., 2013) with the vocational requirement of specialized training for accreditation (Buhai, 2012). As a result, there is growing interest for federal coaching policies to include coaching occupational standards, codes of conduct, qualifications for education and legal recognition, and policies for ongoing vocational development (Duffy et al., 2011). Canada, Australia and the United Kingdom have undergone evaluations of the effectiveness of their formal coach education programs to meet occupational standards (CAC, 2005; MORI, 2004; Woodman, 1984). Their resulting policies have transferred responsibility to national sport

organizations (responsible for individual sport rules) to establish accreditation programs that require competency-based education and training for coaches.

vi. International Trend of Sport Coach Development. Sport is an important cultural and societal experience, as it often unites citizens through pride for the nation's athletes at major games (Comeau, 2012; Heere et al., 2013). In addition, it provides citizens with healthy activities, social interaction and relaxation (Sport Canada, 2012). Knowing this, several international governments have budgeted federal funding for sport participation, development of athletic experiences, and success at international competitions. Although each nation has different governmental policies, many have a recreation or sport agenda linked with cultural identity (Nicholson, Hoyer & Houlihan, 2011).

In several International Council for Coach Education (ICCE) partner nations, sport policy includes coaching development, where the national sport organizations (NSO's) follow set processes to receive government grants (Green, 2007). These policies exist when coaches are valued for providing public participation through positive sporting experiences, and for their contribution to athlete success at all levels (Taylor & Garratt, 2010b). These countries aim to improve the quality of coaching by providing education, training, mentorship and accreditation. Unfortunately, differences in policies and programs for coaching development have resulted from nations working independently. These differences pose a challenge to establishing an international standard coaching career framework that the ICCE seeks (ICCE, 2010).

Federal governments, including the United Kingdom, Australia and Canada, have shown an interest in developing sport coaches through financial and structural commitments. To this end, governments are particularly interested in the return on investment in national coach education programs. In the United Kingdom (UK) coach training and certification programs require coaches to display a minimum level of competency (Cushion et al., 2010). This was a result of governmental requirements for sport organizations to have occupational standards, policies and accountability for coach development (Taylor & Garratt, 2010a). Likewise, in Australia, coaching has been viewed as a vocation that requires competency-based training for coaches to gain practical experience within a formal training framework (Layton, R., 1995). An evaluation of the Coaching Association of Canada's National Coaching Certification Program (NCCP) in the 1990's discovered a need to improve coach education, and initiated the

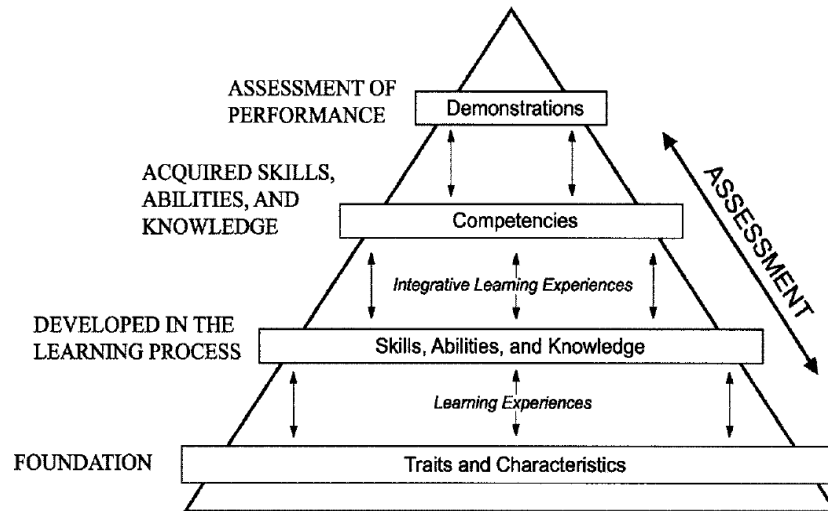
development of coaching competencies through practical experience (CAC, 2005). As a result, changes were implemented to their coach education program to prepare coaches through competency-based education and training (CAC, 2011a).

2.3 Competency-Based Training

i. Competency in Human Resources Management. The growing trend of competency-based training for coaches created questions about what competency is, how it is developed, and how the transition impacted sport organizations (Schembri, 1995). Research from Human Resource Management has developed competency-based models and policies for employers to fill positions in their organizations based on a combination of knowledge, skills, attitudes, and personal attributes such as values and personality (Bowen, Ledford, & Nathan, 1991; Evers, Rush, & Berdrow, 1998; Patterson et al., 2000). Competency is an individual's ability to perform a role or act in an appropriate way, such as for employment (Dubois & Rothwell, 2004; McClelland, 1973). An individual's competency can be seen through their behaviours, as observable actions based on their knowledge, skills, personal characteristics, and aptitude (Lucia & Lepsinger, 1999). Having strengths in only one of these would not be sufficient. To ensure success, individuals are hired if they show competency and fit the organization's culture and values (Bowen et al., 1991; Lucia & Lepsinger, 1999).

Research on competency-based human resource management has produced three meanings of competency (Hoffman, 1999). Firstly, competency denotes an observable performance, or an output of an individual's work. The outputs, or tasks, are used to design competency-based training. Learners are progressively taught how to perform the tasks and are assessed on their performance. A second meaning of the term competency includes standards of performance with multiple levels such as minimal, competent and exceptional. These standards are observed to determine whether requirements are met, or to provide feedback to the employee regarding the quality of their performance. The third meaning of competency credits a person's individual attributes such as their knowledge and attitude as requirements for competent performance. This approach identifies the knowledge, skills and attributes that a competent performer possesses. As shown in Figure 1, the interrelationship of inputs to reach a desired output (performance) is complex. Each level of the pyramid builds a foundation to the next.

This progression is important for educators and learners to understand their roles in the learning process and the acquisition of competencies desired for success in a vocation such as coaching.



(U.S. Department of Education, 2001, p.8)

Figure 1. A performance is individual in nature and based on the interrelationship of a person's traits and characteristics with learned knowledge and skills.

Employees have traditionally been hired based on their knowledge, skills and abilities (KSA's) (Bowen et al., 1991). Hiring an employee relies on a clear description of these KSA's. A job task analysis is common practice to determine both the tasks an employee does and the knowledge and skills needed to perform them (Blank, 1982). The task analysis becomes a detailed job description from which human resources departments hire employees with the appropriate credentials to match the job. Competency-based human resource models are centered on job requirements of qualifications, experience, and capability to show success through the analyzed evidence. As noted above, having competency goes beyond simply showing an acquired skill or knowledge, but rather requires the interrelationship of KSA's with individual traits and characteristics.

As a result of these changing recruitment strategies, new methods for employee training and development are emerging. A common method to prepare trainees to receive accreditations

specific to jobs is referred to as Competency-Based Education and Training (CBET). But, competence is “a complex and sophisticated construct in that it involves a person doing, thinking, reacting, responding and communicating, all within the environment of work and to the standards of performance expected in the workplace” (Gibb, 1993, p.58). The competencies expected of learners are communicated as clearly defined performances that can be observed. It is beneficial for learners to also be aware of variations in performance that would make them successful or unsuccessful (Bowden, 2000). Through CBET individuals are supported to complete the education and training at their own pace, and receive accreditation when the minimum standards for all of the expected outcomes are achieved.

The benefit of CBET for both potential employees and employers is that the training provides a direct link between curriculum and employment. This is the reason it is a method of choice in training in many professional fields. For employers, CBET gives them the ability to recruit candidates with ideal characteristics and traits that match their organization’s vision, mission and organizational objectives (Bowen et al., 1991; Mansfield, 1996). The CBET model has additional benefits to learners because they are at the center of the educational and assessment activities. They are provided with a transparent explanation of the required expectations from the beginning of their training (Sonnadara et al., 2013). It should be clear that both educators and learners are striving for the same goals. In addition, Competency-Based Education and Training is known to be quicker and leads to better performers than traditional knowledge-based programs (Sonnadara et al., 2013).

ii. Competency Assessment. Assessment is a process of collecting evidence to determine if competency has been achieved (Gibb, 1993). Some authors have acknowledged that instruction, learning and assessment must be aligned in order for a CBET program to be successful (Baartman et al., 2006; Hager & Gonczi, 1996). Assessment begins by determining the desired performance. Upon determining performances for competency, activities are designed for education and training for each individual to engage in the learning process (Baartman et al., 2006). In CBET, evidence of competency is collected throughout the learning process as learners are expected to acquire competencies through a combination of education, training, problem solving, application and reflection (Gibb, 1993; Voorhees, 2001).

To measure the impact of competency-based instruction, an analysis of the learner’s

ability is made on specific performance criteria (Leung, 2002). A clear description of the criteria should be published prior to assessment to encourage learners to understand what is expected of them (Ling, 1999). The best competency assessments have criteria that have been agreed upon by experts in the field, making them relevant to trainees for current and future employment (Lichtenberg et al., 2007). Learners and assessor(s) should agree on the purpose of the assessment and reflect on it throughout the learning process (Baartman et al., 2006).

When assessment of competency is used to award an accreditation, such as a vocational or professional license, the results are highly valued by the learner. Performance assessment should be designed with intention to give feedback throughout the learning process (Sonnadara et al., 2013; Voorhees, 2001). A valid assessment is desired as it measures performance criteria and provides accurate feedback about the targeted performance. For this reason, competency assessments are found to be most authentic when they utilize an assessment instrument relevant to the context in which it is applied (Baker, O'Neil, & Linn, 1993). The goal of competency assessment is to use instruments that are valid and reliable, as well as fair and flexible (Booth, 2000; Thomson, Saunders, & Foyster, 2001).

Some methods of assessment have been criticized as being singularly focused on skills. Competence is not simply measured by a checklist of observable behaviours, but rather a holistic view of personal attributes that lead to competent performance. As stated by Hager and Gonczi (1996, p.16): “while performance of tasks is directly observable, abilities or capabilities that underlie the performance are necessarily inferred”. When compared to other forms of assessment, “the competency based approach consists of functional analysis of occupational roles, translation of these roles into outcomes, and assessment of trainees' progress on the basis of their demonstrated performance of these outcomes” (Leung, 2002, p.693). For these reasons, competency assessments require significant effort in the design of appropriate measurements and more time for the collection of evidence (Booth, 2000). Despite the challenges of assessing competency, both learners and employers have expressed the benefits of the approach for its high quality training and predictability of job performance (De Vos, De Haw, & Van der Heaven, 2011; Dubois & Rothwell, 2004).

iii. Competency Assessment of Sports Coaches. The trend in sport coach education to adopt Competency-Based Education and Training (CBET) has gained popularity worldwide. In

Australia and in the United Kingdom, governmental policy has dictated a CBET program for sports coaching (Duffy et al., 2011; Layton, 1994). In both cases, the CBET approach has resulted in a reformat of their coach education programs to include the collection of evidence for competency assessment: “certification of coach learning demonstrates that coaches have satisfied governing bodies of sports’ quality assurance criteria by acquiring and displaying a desired minimum level of competency” (Cushion et al., 2010, p.2).

In Canada, a review of the Coaching Association of Canada’s large coach education program (NCCP) highlighted the need to improve the quality of coaching through improved delivery of coach education and training (CAC, 2005). Changes from this review included a shift from traditional theoretical instruction of coaches to a CBET model (CAC, 2005). This model encourages coaches to not only acquire knowledge but to be more aware of what they do with their knowledge. Educators are hopeful that the quality of coaching will be improved with an emphasis on more experiential learning and assessment of coaching competence in the specific sport context (Demers, Woodburn, & Savard, 2006).

To date, there is a dearth of research on coaching competency assessment. The International Council for Coach Education (ICCE) has suggested that nations who implement a CBET model should provide clear descriptions of competencies (Duffy et al., 2011). While the Coaching Association of Canada administers policies for competency-based coach education and assessment, each sport is entrusted to identify the tasks, activities and competencies of coaches in both recreational and competition contexts. For accreditation to be granted, coaches must demonstrate competence in the required criteria through applications of knowledge and skills in the sport environment. In a CBET model, the education and training focus on preparing the coach to be successful in the assessment. Although substantial research exists on coaching pedagogy and curriculum, there are no known studies investigating how coach education enables coaches to achieve competency, or how positive competency assessments allow coaches to achieve career goals (Hay et al., 2012).

In addition to human resources CBET models, the field of education offers another classification of competency as a stage of professional expertise. Studying how teachers develop has provided insight into coach education because both are pedagogical endeavors (Bergmann Drewe, 2000; Jones, 2006). For example, Berliner’s (1988) stages of expertise in teaching have

been applied to coaching (Schempp et al., 2006). Each stage describes a level relative to an observable performance of the individual. By using descriptions of performance, it is possible to differentiate which stage of development the individual is in. Beginner coaches look for objective facts and features in their lesson but are unable to see relationships between events or between lessons. Competent coaches are still rule oriented but can apply them with respect to the context and situation. They are able to use their experience to strategically solve problems and analyze results. Proficient coaches are analytical yet display logical progression in their instructional decision-making they are in control and can influence athlete changes rather than simply identify results. The mastery stage of expert is descriptive of individuals who are naturally outstanding performers, thinkers, and decision-makers who are fluid, flexible and effective in the coaching environment. According to these stages, a coach may behave consistently as an expert, or may be in transition through earlier stages of Beginner, Competent and Proficient. Although the stages seem linear, the ability to move through stages and how quickly depends on the individual. The presentation of expertise has positively encouraged coach training to shift from assessing knowledge to assessing behaviours (Cassidy & Kidman, 2010).

2.4 Testing Reliability and Validity of CBET Programs

Reliability is the consistency of an instrument to yield the same data over multiple occasions or by multiple observers, while validity is the degree the evidence is used to support actions or interpretations (Kottner et al., 2011; Messick, 1989). Together, reliability and validity provide valuable information about an instrument. A reliable instrument that is not valid is not very useful as it shows repeatability of potentially erroneous data. Furthermore, the validity of an instrument is dependent on its reliability to measure the desired attribute (DeVon et al., 2007). When new competency-based instruments are designed they must be carefully constructed to measure the desired performance outcomes accurately (Dubois & Rothwell, 2004). The use of existing instruments requires assessors to decide if they are appropriate for the situation and construct. Instruments are often assessed and reported on for their strength and utility to allow users determine the trustworthiness of the data they collect (Voorhees, 2001), most commonly reported as validity and reliability. This is important for observing a coach's performance, where

the instrument should confidently present data about the coach in the intended coaching context (Brewer & Jones, 2002).

i. Reliability. Reliability is a measure of confidence that the results of an assessment can be repeated under the same conditions (Thomas, Nelson, & Silverman, 2011). In theory it is desirable for an assessment to be given under the same conditions every time; however, in practice this is not always possible. Changes may occur from one testing scenario to another in any of the following: the subject's mood, the subject's knowledge of the instrument, the environment and potential distractions, and the people administering or rating the observation may differ in their amount of knowledge and experience (Murphy & Davidshofer, 2001). A reliable instrument is desired as it indicates that a consistent result can be found even when these changes in conditions occur.

Reliability is measured through "a quantitative estimate of the amount of measurement error caused by the scoring inconsistency of the raters" (Fan & Chen, 2000, p.532). The more consistent the ratings, the more confidence raters, researchers and practical users have in the instrument (Gwet, 2008). The consistency of raters is analyzed through the similarity of their ratings. This is also known as agreement (Kottner et al., 2011) where perfect agreement is 1.000 (or 100%) and the total absence of agreement is 0.000 (Hayes & Krippendorff, 2007). Landis and Koch first published benchmarks to report the range of agreement (Table 1).

Table 1.

Strength of agreement on a benchmark scale. Benchmarks are commonly reported in the results of agreement and reliability studies (Gwet, 2010; Kottner et al., 2011; Landis & Koch, 1977).

Statistic	Strength of Agreement
<0.0	Poor
0.0 to 0.20	Slight
0.21 to 0.40	Fair
0.41 to 0.60	Moderate
0.61 to 0.80	Substantial
0.81 to 1.00	Almost Perfect

When different raters use the same instrument to assess the same event, it is valuable to know the consistency between them. This is known as inter-rater reliability (Hallgreen, 2012; Kottner et al., 2011). To ensure reliability analysis is accurate, raters must provide their ratings independently. Independence requires raters to observe the event and provide their ratings on their own without knowing what other raters have given (Mellenbergh, G., 2011).

When the same rater provides ratings of the same event multiple times, reliability can also be measured. The consistency of ratings by a single rater over multiple observations is known as intra-rater reliability (Kottner et al., 2011). In this case, independence of ratings is controlled by a rest period between observations to allow memory lapse (Brewer & Jones, 2002).

The ratings on an instrument are used to make an inference about a subject or an event (Kundel & Polansky, 2003). For example, assessment results are used to determine if a candidate should be awarded a vocational or professional certification. When agreement exists, the assumption is that the raters have measured relevant information on which to make a decision. In order to determine if consistent ratings are also accurate, a measure of the test's validity is needed.

ii. Validity. Validity is the degree to which ratings are appropriately used to make decisions (Messick, 1989). It is important for all forms assessment, including observation and rating of performance (Baker et al., 1993). High validity confirms the instrument is capable of measuring the attribute it is designed for and can be determined either by the relevance of the items, or by the ability of the instrument as a whole to rate an attribute (Polit & Beck, 2006).

There are two types of validity that specifically apply to a behaviour analysis for competency assessment: content validity and face validity. Content validity is the extent to which an instrument measures what it is intended to measure (van der Mars, 1989). It is validating whether items adequately represent the attribute for which it was designed. The most common method used to examine this type of validity is a content validity index (CVI). In the administration of a CVI, the instrument or individual test items is given a quantifiable rating by context experts (Lynn, 1986; Polit & Beck, 2006; van der Mars, 1989). Polit and Beck (2006) distinguish two types of content validity indices. The first is an item level content validity, I-CVI where the expert's I-CVI ratings are used to "guide decisions about item revisions or

rejections, and the experts' comments would guide the development of any new items" (Polit, Beck, & Owen, 2007, p.466). The second CVI is a single rating for the entire instrument, where a statistical analysis is based on the proportion of items reported as relevant on the instrument. Unfortunately in this second method an average of the I-CVI's is given without providing item specific values or resulting actions for the low scoring items. For this reason, researchers are encouraged to report I-CVI ratings and report how that guided revisions to the instrument (Polit & Beck, 2006).

The second type of validity is face validity: the relevance to the group for whom it was designed and the environment in which it should be used (Furr & Bacharach, 2008; Thomas, Hathaway, & Arheart, 1992). This process signifies how relevant an instrument is, or the items are, to the competencies performed by the individuals being assessed. This may include the viewpoints of non-experts who are experienced with the competencies and can compare what they do with what is expected to pass the assessment.

Unlike content validity, which is determined by experts, face validity relies on people who hold practical interests in the assessment process. Nevo (1985) explains that the following groups should be surveyed:

- a. The persons who actually take the test (e.g., job applicants, participants in experiments, school pupils, etc.);
- b. The nonprofessional users who work with the results of the test (e.g., personnel administrators, employers, admissions officers, chairpersons of university departments, psychiatrists, etc.); and
- c. The general public (e.g., newspaper readers, newspaper reporters, parents of testees, judges, politicians, etc.)" (p. 288)

Face validity is a non-statistical estimate of the logical tie between items on an instrument and its purpose (Lynn, 1986). How researchers collect this information varies between quantitative ratings and qualitative surveys. Quantitative data would be reported in similar ways as content validity (Nevo, 1985), while qualitative data from the group surveyed would be interpreted to provide clarity and credibility for the instrument (Baartman, et al., 2006).

Together, content validity and face validity can provide two different valuable pieces of information about an instrument. The two groups surveyed can identify the relevance of test items as well as add items that are relevant but missing. Both forms of validity are essential for

instrument design and acceptance especially when used for observations of performance (Baker et al., 1993). They ensure the instrument is trustworthy, accurate, and suitable of the time and effort of both assessors and learners (Linn, Baker, & Dunbar, 1991). Select instruments should be used within an assessment to improve instruction and learning (Baartman et al., 2006). Therefore, validity of competency assessment instruments is essential for the learner to receive accurate performance feedback (Booth, 2000; Hager & Gonczi, 1996; Thomson et al., 2001).

2.5 Purpose of the Study

Formal coach education programs, including the Coaching Association of Canada's (CAC) National Coaching Certification Program (NCCP) (CAC, 2005), have adopted Competency-Based Education and Training (CBET). Gilbert and Trudel (1999) showed that the CAC's older NCCP, as a knowledge-based source of coach education, was ineffective at providing new knowledge or changes in coaching behaviours. Changes in the NCCP were made to include an assessment for evidence of coaching competency in context specific accreditation, but limited research has been done to determine if the competency-based approach is more effective than the knowledge-based approach. The purpose of this study is to follow a previously published five-stage process to establish validity and reliability of a competency assessment for coaching certification.

Hay et al. (2012) highlight the urgent need for greater understanding of assessment practices in the field of sports coaching. It is recommended that organizations report the accuracy and consistency of their CBET assessments (Booth, 2000; Thomson et al., 2001; van der Vleuten et al., 2010). Despite a significant body of literature on sports coach development and the use of CBET, no known research is available regarding the quality of coaching competency assessments (Cushion et al., 2010). CAC refers to their pre-competition levels (recreational context) as Community Sport. In this context, athletes learn basic skills in a safe and fun environment ("New NCCP Generic Overview" n.d.). Coaches that deliver these experiences are trained through context and sport specific NCCP courses. To receive certification, coaches must complete the training and assessment components as required by their national sport organization.

3. METHODS & RESULTS

3.1 Instrument

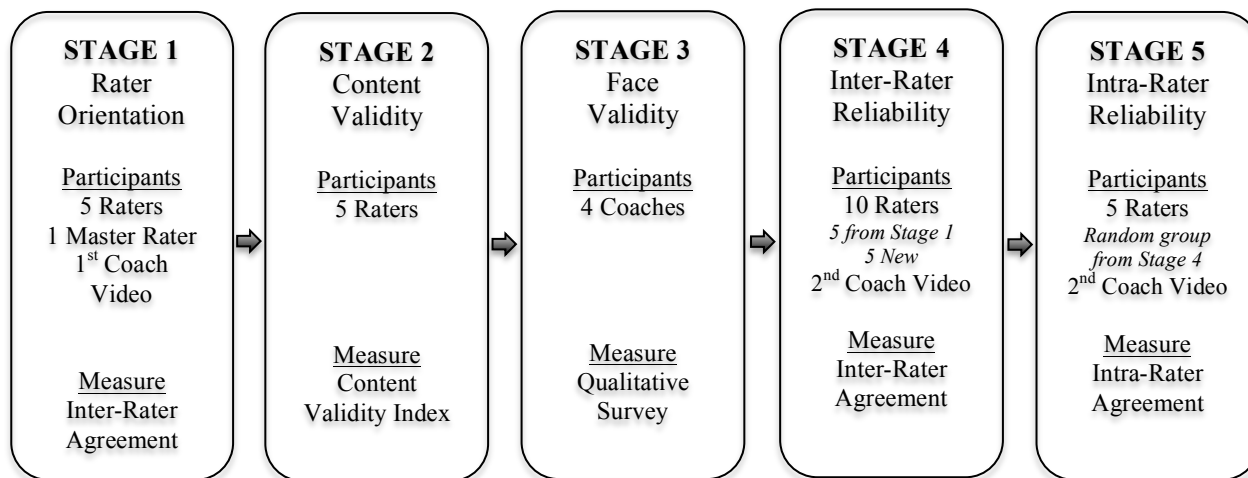
In collaboration with Gymnastics Canada, the assessment instrument for Community Sport artistic gymnastics coaches, The Coach Evaluation (Gymnastics Canada, 2011), was tested for validity and reliability. The instrument was developed by Gymnastics Canada with assistance from Coaching Association of Canada (CAC) and Gymnastics Australia. The instrument included fifty-four items on which coaches were assessed for competency. The items were separated into sections in the order they would occur: before the lesson (four items); within the warm-up (ten items); within the lesson (16 items); within the cool down and conclusion (six items); and general coaching throughout (18 items). The items were written as questions for the rater to answer on a dichotomous scale of *Yes* or *No*. For this study, the instrument was transferred from the original Portable Document Format (PDF) version to a Microsoft Word electronic rating form.

3.2 Five Stage Process

Brewer and Jones (2002) stress the importance of testing context-specific validity and reliability of coaching observations. Their methodological process measures content validity, face validity, inter-rater reliability and intra-rater reliability. Their process follows five stages (Figure 2) and has been used in several studies to establish validity and reliability of observation instruments for coaches and physical education teachers (Cushion, Harvey, Muir, & Nelson, 2012; Ford, Yates, & Williams, 2010; Roberts & Fairclough, 2012). Ethical approval was exempt for this study as it was based on program evaluation.

Two groups were needed to complete the five stages, a group of experts (raters) and a group of coaches. The participants for each stage are displayed in Figure 2. The procedures and results for each stage will be presented below as they were completed, in sequential order.

Figure 2. A five-stage process for establishing validity and reliability of a gymnastics coach observation instrument (Brewer & Jones, 2002).



3.3 Stage 1 Rater Orientation

i. Stage 1 Procedures. The objectives of the first stage included an orientation to the instrument, practice using the video observation system, and practice rating while observing a coach (van der Mars, 1989). This required five raters, one master rater and a video of one coach's lesson.

The raters were experts in Community Sport Artistic Gymnastics as they were NCCP Learning Facilitators trained by Gymnastics Canada to deliver the education programs for new coaches in the Community Sport (introductory) context of gymnastics. To become Learning Facilitators, individuals were nominated by their province based on their education and experience in the context. At the time of their nomination, they were at least 18 years of age and certified NCCP Level 2 in Artistic Gymnastics. Individual background information on coaching experience, or number of years as a Learning Facilitator was not collected. Gymnastics Canada provided approval and email addresses of Learning Facilitators. All twenty-six potential raters received a letter by email (Appendix A). A total of ten agreed to participate: four males and six females. The ten raters were from several provinces including two from British Columbia, two from Saskatchewan, three from Ontario and one from each of Quebec, New Brunswick and Nova Scotia. For subsequent stages, all ten raters were needed, but for Stage 1, only five of the ten

were required. Five were randomly chosen to an orientation to the instrument and the rating process. These five raters included two males and three females.

The master rater was an NCCP Master Learning Facilitator conveniently located in Saskatchewan, and was the author of the coach education textbook for Community Sport Artistic Gymnastics. Master Learning Facilitators train, evaluate and mentor Learning Facilitators after demonstrating commitment, excellence and many years of practice in the coaching context.

The coaches to be assessed were required to be working in Community Sport Artistic Gymnastics. They had a minimum of five years of coaching experience and Level 1 NCCP certification (the nomenclature used prior to the introduction of the CBET model). Recruitment of four coaches, three females and one male, occurred through emailed letters from Gymnastics Saskatchewan to their member clubs and coaches (Appendix B). Four coaches from Saskatoon and surrounding region agreed to have a regularly scheduled gymnastics lesson video recorded in their own facility. Since only two videos were required to complete the research, the best two of the four coaches' videos were chosen based on the quality of the picture and sound, setup of the gymnasium, and whether the overall coaching reflected the Community Sport context. The sample of four coaches was required for a latter stage as the appropriate panel size for face validity testing (Brewer & Jones, 2002).

Before video recording, all gymnasts submitted parental consent to be video taped for this research. The researcher, with an Apple iPad, collected the video, and the audio was collected through a compatible Bluetooth microphone. The microphone was connected to the iPad through the freedomMIC application (Scosche, 2012). The collected videos were downloaded and edited in iMovie. To reduce time required to send mail videos and eliminate concerns that downloading electronic video files would not be possible for all raters, the observations took place through online meetings via <http://www.webex.com>. The raters were required to have an Internet connection to access the meeting, and an email address to receive the observation rating forms. The video selected for Stage 1 was of a female coach working with six girls between six and seven years of age practicing in Community Sport gymnastics.

The observation process occurred at a meeting time agreed to by the rater and the researcher. The following process was used (detailed instructions that the raters received are enclosed in Appendix D):

- raters were emailed a unique meeting number that could only be accessed at the meeting time
- an electronic version of the rating form was sent in advance by email
- raters were asked to read and review the instrument in advance
- at the meeting time, raters entered the meeting by the email link
- each rater completed the observation independently and at different times
- the researcher led the session and introduced the rater to the observation session
- the researcher ensured the rater had audio capability and the rating form present
- the video was played and controlled by the researcher
- the coach's lesson was approximately one hour
- the rater observed the entire video of the coach and completed the ratings while observing
- raters used electronic or paper copies of the rating form
- after the observation, the online meeting space was closed and raters emailed their ratings to the researcher

The data collected in this stage allowed reliability of the instrument to be determined. The inter-rater reliability between raters and the master rater was calculated as was agreement between the five raters.

The most common method to measure rater reliability is Percent of Agreement (Formula 1). This method was used for inter-rater reliability in Stages 1 and 5, and intra-rater reliability in Stage 4. The calculation entails counting the number of times raters agreed out of the total number of ratings. The greater the percentage found, the greater the reliability.

$$\text{Percent of Agreement} = \left[\frac{\# \text{ of Agreements}}{\text{Total \# of Observations}} \right] \times 100 \quad (1)$$

(van der Mars, 1989; Hallgreen, 2012)

The one major concern with Percent of Agreement is that it does not account for agreement by chance (Fleiss, 1971; Hallgreen, 2012;). The impact of not accounting for chance is an overestimated reliability statistic, which could lead to improper inferences about the participant (Watkins & Pacheco, 2000).

To alleviate this overestimation, Gwet's Agreement Coefficient (AC_1) (2002a; 2002b; 2010) was chosen to calculate rater reliability (Formula 2). AC_1 is calculated by first determining the Percent of Agreement, where P_o is the number of agreements divided by the total number of observations, P_e is the probability of chance agreement, and P represents the

approximate chance that a rater (A or B) selects category 1. Although AC_1 is fairly new, it is the most robust statistical measure of agreement (Gwet, 2002) and has gained popularity as a method of choice in studies measuring rater reliability (Wongpakaran et al., 2013).

$$AC_1 = \frac{Po - Pe(\gamma)}{1 - Pe(\gamma)} \quad (2)$$

$$Pe(\gamma) = 2P(1 - P) \quad P = \frac{(A1+B1) \div 2}{N}$$

(Gwet, 2002a)

The calculations for Percent of Agreement and AC_1 were completed with AgreeStat 2011.2 for Mac (Advanced Analytics, 2011) to measure rater reliability in Stages 1, 4 and 5. For both calculation methods, the resulting reliability statistics were reported as a value between -1.0 and 1.0 and compared to the benchmarks from Table 1 to report the strength of the agreement (Gwet, 2010; Kottner et al., 2011; Landis & Koch, 1977).

ii. Stage 1 Results. The ratings for each rater on 48 items were compared to the master rater's by analysis of inter-rater reliability as shown in Table 2. Four of the raters agreed with the master rater at *Substantial* or *Almost Perfect* levels, while one expert mostly disagreed with the master rater, with a level of agreement at *Fair*.

Table 2.

Stage 1 Results. The Levels of Agreement Between Five Raters and a Master Learning Facilitator.

Rater	Agreement Coefficient (AC_1)		Percent of Agreement
	Coefficient	Benchmark	
Rater 1	.73	Substantial	.79
Rater 2	.24	Fair	.54
Rater 3	.95	Almost Perfect	.96
Rater 4	.67	Substantial	.75
Rater 5	.91	Almost Perfect	.92

Note. Ratings were provided for 48 items in the observation of one coach during a one-hour lesson. The agreement is based on the ratings by all five raters and the master learning facilitator using the Gymnastics Canada coach assessment (Appendix D).

3.4 Stage 2 Content Validity Testing

i. Stage 2 Procedures. In the second stage, raters were asked to judge the relevance of each item on the instrument by completing a Content Validity Index (CVI). The objective was to determine whether the items listed in the instrument were essential to the underlying construct (Lynn, 1986), (coaching Community Sport Artistic Gymnastics). A minimum of five rater opinions is recommended for content validity testing (Lynn, 1986), and all five raters from Stage 1 completed the CVI.

The CVI was distributed via email to the raters as a Microsoft Word macro-enabled document (Appendix E). The raters were given a detailed description of the context the instrument was designed for and explicit instructions on how to complete the CVI. The survey included rating boxes for each item and text space to include comments when they rated an item as non-essential. In addition, raters were asked to add any items that were essential to coaching in the Community Sport context but were not included.

Following procedures by Lynn (1986) and Polit and Beck (2006), a CVI for all fifty-four items in the instrument was calculated. Each rater completed the CVI independently. The ratings were on a Likert scale from 1 (*the behaviour is not essential*) to 4 (*the behaviour is very essential*). Ratings on the upper half of the scale indicated that the item was essential for a coach to perform in the context. A four-point scale was selected, as it did not allow respondents to choose a neutral midpoint.

For each item, the ratings of 3 (*quite essential*) or 4 (*very essential*) were used to determine the item level CVI. With only five raters, the minimum I-CVI is 100%. Therefore, all raters had to rate the item 3 or 4 for it to be considered valid (Lynn, 1986; Polit & Beck, 2006). If the item received a single rating of 1 (*not essential*) or 2 (*somewhat essential*) it was removed from the instrument.

Raters were asked to comment why they rated an item as 1 or 2 and if a rewording would improve the rating of that item. Raters were also asked if there were any items not included that are essential to coaching Community Sport Artistic Gymnastics (Lynn, 1986). The researcher read the responses and modified the instrument for every suggestion, except when the rater had suggested an item that was already included but described in a different way. In these cases, the wording of the item was reviewed and modified to match the preference of the rater.

ii. Stage 2 Results. From the CVI calculations (Table 3), five items were removed for receiving one rating below 3 (*quite essential*), $I-CVI = 0.8$. One item was removed for receiving two ratings below 3 (*quite essential*), $I-CVI = 0.6$.

Table 3.

The Number of Expert Ratings by Item.

Item	Not Essential	Somewhat Essential	Quite Essential	Very Essential	I-CVI (%)
Before the lesson, did the coach...					
Arrive a bit early with a written lesson plan?			2	3	1.0
Have equipment set-up and ready for use for the lesson?			1	4	1.0
Inspect the equipment for safety?		1		4	0.8*
Greet each participant as he/she arrived?			1	4	1.0
Within the warm-up, did the coach...					
Assemble participants appropriately to start the warm-up?			3	2	1.0
Conduct a balanced warm-up (were all major muscle groups worked)?			2	3	1.0
Did the warm-up start with energetic activities (fast and fun moving games)?			1	4	1.0
Incorporate physical preparation (ESPF) activities?			3	2	1.0
Use activities appropriate to the participants' age/abilities?			2	3	1.0
Involve all the participants?			1	4	1.0
Increase the participants' body temperature/heart rate appropriately?			2	3	1.0
Maintain good class control?				5	1.0
Use music (if applicable)?	1	1	3		0.6*
Move participants onto the 1st activity (main part) quickly and appropriately?			3	2	1.0
Within the main part of the lesson, did the coach...					
Incorporate physical preparation (ESPF) and motor preparation activities?			1	4	1.0

Break skill learning into sequential steps?		1	4	1.0
Use drills / progressions / activities to the participants' age/abilities?			5	1.0
Use different teaching methods for participants with different learning styles?	1		4	0.8*
Use a variety of drills / progressions / activities to ensure lots of little successes?			5	1.0
Recognize and adapt for an individual's needs?			5	1.0
Stress 1 to 3 key coaching points and safety?	1	1	3	0.8*
Provide clear and concise instructions?		2	3	1.0
Correct errors as they occur by providing constructive & specific feedback?		1	4	1.0
Provide feedback to reinforce effort, trying, etc.?		1	4	1.0
Have everyone involved through continuous activity?			5	1.0
Supervise all participants at all times?			5	1.0
Use supplementary equipment to assist with teaching/ learning of skills?		3	2	1.0
Use equipment (large and small) effectively?		4	1	1.0
Do activities that work towards / achieve the stated objectives of the lesson?		1	4	1.0
Maintain class control?			5	1.0
Within the cool-down, did the coach...				
Assemble participants appropriately to start the cool down?		2	3	1.0
Involve all participants?			5	1.0
Maintain good class control		1	4	1.0
Do slower paced and/or stretching activities?	1	2	2	0.8*
Finish off with a fun activity?		2	3	1.0
Within the conclusion, did the coach...				
Organize the participants appropriately to leave the gym?	1	3	1	0.8*
Overall communication: did the coach...				
Catch the participants' attention quickly?			5	1.0
Provide clear and concise feedback and instructions?		2	3	1.0
Check the participants' understanding with regards to instruction and safety?		2	3	1.0
Maintain good, clear voice control?		2	3	1.0
Use non-verbal communication with the participants?		3	2	1.0
Address participants respectfully?			5	1.0
Have fun with the participants?		2	3	1.0
Handle disciplinary situations appropriately?		1	4	1.0
Handle any injury / emergency situation that occurred appropriately?		1	4	1.0
Dress appropriately for the lesson?		4	1	1.0
Act in a professional manner at all times?		1	4	1.0
Show enthusiasm?		2	3	1.0
Display a positive attitude towards their coaching?			5	1.0
Present the right amount of material?		3	2	1.0

Present material in a logical way?	3	2	1.0
Present material at the appropriate level for the participants?		5	1.0
Demonstrate a sound knowledge of the material being presented?	2	3	1.0
Cover all key points / information relevant to the desired learning experience?	2	3	1.0

Note. *Items <1.0 *I-CVI* were removed from the instrument (Polit & Beck, 2006).

As stated by Brewer and Jones (2002, p. 146), “content validity...will denote that the instrument completely records or adequately samples the principal range of behaviors that are demonstrated by coaches within the specific environment”. Raters expressed that items essential to coaching in the specific sport context should be added to the instrument by answering the question: (a) *Are any behaviours omitted that are essential to coaching Artistic Gymnastics in NCCP Community Sport?*

Based on expert input, four existing items were modified:

- Original item: Address participants respectfully?
→ Modified: Address participants with respect and kindness?
- Original item: Have fun with the participants?
→ Modified: Have fun with the participants (ex.use humour)?
- Original items: Have equipment set-up and ready for use for the lesson?; Inspect the equipment for safety?
→ Modified: Setup equipment to be available, ready and safe for use in the lesson?
- Original items: Break skill learning into sequential steps?; Use drills/progressions/activities appropriate to the participant’s age/abilities?
→ Modified: Teach skills by sequential steps (drills/progressions)?

And one new item was added:

- End the class on a positive note

In the instrument studied, items are competencies that are essential for coaches in the sport context, as verified by the group of raters. Competencies are specific outcomes that are essential for success in the occupation (Blank, 1982). Non-essential items were removed and additional items were modified or added. From these changes, the instrument’s quality and

relevance are affirmed (Lynn, 1986). Because few changes were made, there was no further need to review the items or the underlying construct being measured (Polit & Beck, 2006).

In Stage 2, a second question was asked of raters regarding the design of the instrument: (b) *Does the scoring rubric (YES/NO selection) appropriately measure the behaviours? Why or why not?* Responses by five raters are shown in Table 4. The expert group expressed that the scale is not sufficient to rate the range of behaviours possible within the items. For this reason they preferred a range of ratings to accommodate the range of performances.

Table 4.

Stage 2 Responses to Content Validity Survey Question (b) by Five Raters.

Rater	Are any essential items omitted from the assessment instrument?	Rationale
1	Yes	All behaviours exist or do not, for many of them there are ranges to be expanded on.
2	No	Yes or no is often too drastic. The level of competency with a kind of scale is probably better. Yes or no may be correct for the question: "Does it meet the minimum requirement for the task?"
3	No	I found it very difficult to judge YES or NO for some of these points in the video. I would rather see a rating of 1-4 (1 being behavior not present, 4 being excellent), or something similar, with some space to provide feedback. I believe that even an evaluation is a time to give a coach some feedback and suggestions, especially given that not all coaches will pass the evaluation the first time.
4	No	Coaches sometimes meet the criteria and I do not feel it is right to say no. Some criteria have more than a yes or no answer.
5	No	I do not feel that yes or no rating is effective, perhaps a below expectations, meets expectations, and exceeds expectations would better service a well-rounded evaluation. Secondly, it has never been communicated how many no's equate a fail. Are there any no's that are weighted more (ex. safety?). Also, in this scoring system, one cannot account for strengths or weakness coaching individual apparatus.

A dichotomous scale, such as the *Yes* or *No* option in the original instrument, is faster to complete than a Likert scale. However, the simple scale may not be a valid measure of increasing competence (Hodges, et al., 1999). It is known that an effective assessment of competency requires significant time and effort in which to collect evidence of the learner's abilities (Booth, 2000). Therefore, a brief assessment, or instruments completed in the least amount of time is not favourable. In addition, learners who desire detailed feedback of their performance (e.g., improvements needed to reach a level of performance) cannot be given this through a dichotomous scale. Using stages of development "would yield a more valid and realistic assessment of cognitive structures and processes; one that, ultimately, may be more useful to practitioners" (Karelitz, 2008, p. 269).

To decide if changing from a dichotomous scale to a Likert scale was appropriate, it was important to understand the purpose of the assessment. The assessment should clearly state what the learner is to achieve (Moskal & Leydens, 2000). In the sports coaching context, the Gymnastics Canada assessment was designed to verify that the coach was competent to "develop, implement and evaluate a safe, and both age-appropriate and level-appropriate gymnastics lesson" (Gymnastics Canada, 2010). At the completion of an observation, experts (raters) should be comfortable in their rating of the performance (CAC, 2011b). The responses from the Content Validity Survey, question (b), demonstrated that the requirements for a coach to pass the assessment were not clearly stated (Table 4). Considering the preference of the raters to use a range of performance, a Likert scale was chosen to provide an overall competency assessment for raters. The researcher added four possible ratings of competency to the instrument of Beginner, Competent, Proficient, or Expert based on the stages of teaching expertise (Schempp et al., 2006).

3.5 Stage 3 Face Validity Testing

i. Stage 3 Procedures. The aim of the third stage was to establish Face Validity of the modified instrument. This stage involved the population for which the instrument was designed, that is, the four coaches. Three coaches working in Community Sport Artistic Gymnastics completed the survey. One recruited coach did not respond to the survey request. They were asked to review the modified instrument and provide responses to the following three questions: (a) Are all items reflective of coaching behaviours in the appropriate context? (b) Are important

items omitted? (c) Are unimportant items incorrectly included? The responses determined if the instrument was representative of coaching Community Sport gymnastics.

ii. Stage 3 Results. It is obvious from the responses of the coaches (Table 5) that there were no concerns. This procedure ensured that the assessment items reflected duties and job tasks that experienced individuals performed in the coaching context. As a result, no further modifications to the instrument were needed.

Table 5.

Stage 3 Qualitative Responses to a Face Validity Survey of the Modified Instrument.

Question	Coach 1	Coach 2	Coach 3
1.Are all elements of the instrument reflective of coaching in the appropriate context (artistic gymnastics/ community sport)?	I believe all elements within the evaluation are accurate.	Yes I believe you have all of the necessary elements	Yes
2.Are important elements omitted? If so, which ones?	No	No	No
3.Are unimportant elements incorrectly included? If so, which ones?	I don't think there are any elements that are incorrectly included.	No	No

3.6 Stage 4 Inter-Rater Reliability

i. Stage 4 Procedures. The fourth stage involved all ten recruited raters, the five from Stage 1 who received orientation and participated in content validity testing, and five new raters. The same observation process was followed as in Stage 1 where raters were contacted and provided online meeting logins by email (Appendix H). The modified instrument resulting from Stages 2 and 3 with forty-seven items on a polytomous rating scale was circulated by email to each of the

raters for their observation (Appendix I). All ten raters submitted ratings, two raters missed recording one item each.

The video observed was of a different coach than the one observed in Stage 1. The video was a boys Community Sport gymnastics lesson with six participants led by a male coach. The observation took approximately sixty minutes. The raters used the Likert rating scale to assess the item as “*Excellent*”, “*Good*”, “*Fair*”, or “*Insufficient*”. They also rated the coach on the overall competency rating, selecting between “*Expert*”, “*Proficient*”, “*Competent*”, or “*Beginner*”. The reliability was analyzed as agreement between each of the ten raters, and between the orientation group of raters with the new group of raters.

Due to the changes in the rating scale from a dichotomous scale to a Likert scale in Stage 2, the methods of Agreement Coefficient and Percent of Agreement were applicable; however the analysis required the use of a linear weight to measure accurate agreement statistics on a four-point scale (Warrens, 2013). A linear weight was applied through the AgreeStat statistical program to acknowledge that certain disagreements are more serious than others. The linear weight recognized that raters partially agreed when providing ratings close to each other on the scale (Gwet, 2011). The weight decreased as the ratings became further apart (Table 6).

Table 6.
Linear Weights of Inter-Rater Reliability Coefficients.

	Excellent	Good	Fair	Insufficient
Excellent	1	.667	.333	0
Good	.667	1	.667	.333
Fair	.333	.667	1	.667
Poor	0	.333	.667	1

ii. Stage 4 Results. The resulting inter-rater reliability between ten raters was calculated at *Moderate* ($AC_1 = 0.44$, 95% CI [0.36, 0.51]) and *Substantial* (*Percent of Agreement* = 0.72, 95% CI [0.70, 0.75]) as shown in Table 7. No differences in the ratings were found between the orientation group and the non-orientation group. The level of agreement for the orientation

group of raters was found to be *Moderate* for AC_1 (0.58, 95% CI [0.50, 0.66]) and *Substantial* for *Percent of Agreement* (0.76, 95% CI [0.73, 0.80]). The same levels of agreement between the five raters in the non-orientation group were found: *Moderate* for AC_1 (0.49, 95% CI [0.40, 0.58]) and *Substantial* for *Percent of Agreement* (0.75, 95% CI [0.72, 0.78]).

Table 7.
Stage 4 Results. Inter-Rater Reliability of Ten Raters.

Method	Coefficient	Std.Err.	95% C.I.	p-Value
AC_1	0.435	0.0368	0.361 to 0.51	0
Percent of Agreement	0.722	0.0129	0.696 to 0.748	0

Note. Ratings were given within four possible rating categories ranging from insufficient evidence to excellent evidence to rate one coach's lesson on 47 items of coaching competency.

In addition to rating the coach on each item, raters were asked to provide an overall assessment of the coach's competency (Appendix H). Nine of the ten raters assessed the coach as "*Competent*", "*Proficient*" or "*Expert*", while one of the raters assessed this coach as "*Incomplete*" due to issues concerning the safety of the athletes (Table 8).

Table 8.
Results of Coaching Expertise Assessment by Ten Raters.

	Expert	Proficient	Competent	Incomplete
Number of Raters	2	3	4	1

3.7 Stage 5 Intra-Rater Reliability

i. Stage 5 Procedures. To answer whether individual raters were consistent over multiple observations, ratings were collected for a repeat observation of the same coach and same lesson that was viewed in stage four. Observational procedures from Stage 4 were repeated (Appendix H) a minimum of one week after the first observation.

A group of five raters, randomly selected from Stage 4, completed the second observation independently and without knowledge of results from the first observation. Each rater was measured for their personal intra-rater reliability. The ratings from Stage 4 were compared to new set of ratings to determine intra-rater reliability.

ii. Stage 5 Results. Intra-rater reliability was recorded at levels of agreement between “*Moderate*” to “*Almost Perfect*” (Table 9). Two of the five raters were largely consistent between their two sets of ratings. Another two of the five raters had adequate agreement for the repeat tests, while one of the raters had more noticeable challenges with consistency reporting only “*Moderate*” agreement. The findings show that raters were generally able to repeat ratings of the items in a performance seen on two different occasions. The findings also suggest the raters can apply ratings to the performance of coaching competencies consistently.

Table 9.

Stage 5 Results. Intra-Rater Reliability for Five Raters.

Rater	Agreement Coefficient (AC ₁)		Percent of Agreement
	Coefficient	Benchmark	
Rater 1	0.45	Moderate	0.67
Rater 2	0.75	Substantial	0.81
Rater 6	0.77	Substantial	0.89
Rater 8	0.87	Almost Perfect	0.93
Rater 9	0.82	Almost Perfect	0.90

Note. Agreement is measured between two observations of the same coach using the same instrument after a time lapse.

The five raters provided similar views of the coach’s overall ability when comparing the first to second observations (Table 10). Without rigorous training, the raters showed a capacity to repeat their individual views and interpretations of the coach’s competency.

Table 10.

Competency Assessment Results by Five Raters Over Repeat Observations.

Rater	Observation 1	Observation 2
Rater 1	Expert	Proficient
Rater 2	Expert	Expert
Rater 6	Incomplete	Incomplete
Rater 8	Competent	Proficient
Rater 9	Competent	Competent

One rater who believed the coach did not provide enough evidence to meet the minimum standard of Competent in Stage 4 was the same rater who reported this lack of competency in the second observation. The remaining four raters who saw the coach as Competent, Proficient, or Expert maintained this minimum rating, although two raters differed on their second rating by changing one level above or below their first (i.e., Rater 1 provided Expert for observation 1 but only Proficient for observation 2).

4. DISCUSSION

There is limited research on the impact of implementing a Competency-Based Education and Training (CBET) model for sports coaches despite the international trend for sports organizations to use CBET. This study focused on the assessment portion of the CBET model with the primary purpose to test the validity and reliability of an existing coaching competency assessment instrument. Each stage of the five-stage process resulted in valuable insights into the coaching assessment instrument used for the Community Sport context of Artistic Gymnastics. As a result of modifications made through the stages of research, the instrument is valid to assess competencies, in a moderately reliable manner. The implications of the findings in each stage are discussed below with procedures and suggestions for Gymnastics Canada and other sport organizations to follow for effective implementation of coaching competency assessments.

4.1 Stage 1

Inter-rater reliability between five raters is *Moderate* (*Agreement Coefficient* [AC_1] = 0.56, 95% CI [0.42, 0.69]) and *Substantial* (*Percent of Agreement* = 0.70, 95% CI [0.63, 0.77]). This level of agreement shows that despite expertise in Community Sport Artistic Gymnastics, the raters varied in their view of the item completed. These variances may be due to personal preferences, regional preferences, differences in education and experience by raters, or subjectivity of the items (Murphy & Davidshofer, 2001). One rater commented on the challenge in determining a “Yes” rating despite the coach showing a range of performance.

Following the orientation in Stage 1 it was concluded that the instrument’s reliability is moderately strong. The results showed that error may be present in the assessment of coaching competency. This would decrease the confidence held for the instrument by the organization, the raters and the coaches. These results also tentatively indicated that the instrument needed improvements of reliability. It is unknown in this stage if the results are due to the quality of the instrument or differences in opinions regarding competencies.

To address rater reliability, a standardized rater training session is often conducted with an a priori benchmark. For example, a minimum of 0.80 Percent of Agreement between the rater and master rater (Hallgreen, 2012; van der Mars, 1989) is used to ensure all raters achieve the same standard. If a rater fell below the minimum standard they received feedback on which

areas they rated different from the master rater. This learning effect is known to result in stronger reliability (Tracy et al., 1997).

In this study, a standardized training did not occur. Since the instrument had no reliability testing in its design, nor had it been used for coaching assessments, the master rater was as equally unfamiliar with the rating instrument as was the group of expert raters. If Gymnastics Canada had conducted reliability testing during the selection of the assessment items, a baseline level of agreement for a standardized training could have been included in this stage.

4.2 Stage 2

i. Validity. Messick (1989) presented the argument that assessment instruments for certification in a professional field must be valid, while Hay et al. (2012) added that sports coaching assessments are valid when they reflect knowledge and skills that are essential to the sport and level of athletes. This study used two steps to test the validity of the instrument: experts determined the content validity (Stage 2), and coaches confirmed the face validity (Stage 3).

ii. Content Validity. Content validity was tested through an item-level Content Validity Index (I-CVI). This was done to determine the strength of each item in the assessment of competency. An overall CVI was not asked for, as it would have provided average ratings instead of item-specific data. As a result of the I-CVI tabulation (Table 3), the instrument was revised as follows: 54 items reduced to 47 items; the rating scale expanded from two performance categories to four; and an overall assessment of competency was added. As a result, the modified instrument will measure what it is intended to: that the coach can “develop, implement and evaluate a safe, age and level-appropriate gymnastics lesson” (Gymnastics Canada, 2010).

Following Lynn’s (1986) procedures, modifications to the instrument were made if one of the raters determined an item to be non-essential, or suggested an item be added. Six items were thus removed. Four items that were removed received two or more non-essential ratings, but two items could be questioned as they were rated as very essential by four raters and non-essential by only one rater.

Since four raters viewed the item “Inspect the equipment for safety” as very essential, the value of the item was considered. This led to discussions with Gymnastics Canada staff, and the importance to retain an assessment of a coach’s safety in the lesson by adding the word “safe” to another item that was rated as essential. The resulting item was: “Setup equipment to be available, ready and safe for use in the lesson?”

The second item with one rating of non-essential and four ratings of essential was an assessment of the coach’s ability to “Use different teaching methods for participants with different learning styles”. The assessment of this competency is duplicated in the assessment of the coach’s competency to “Recognize and adapt for an individual’s needs” and was therefore removed.

Performing an analysis of CVI is straightforward; however, providing an explanation where there is a large discrepancy between the ratings is not. This is a concern given there are no procedural guidelines of how to act once ratings and comments are collected. According to Polit, Beck, and Owen (2007), the discrepancy “could mean that the directions to the experts were inadequate, or that the experts themselves were biased, erratic, or not sufficiently proficient” (p. 466). Despite this concern with the CVI procedure, the analysis portrays the instrument’s relevance to the vocation of gymnastics coaching, and allows Gymnastics Canada the opportunity to report the instrument’s validity to potential coaching assessment candidates (Kaslow, et al., 2007).

iii. Assessment Scales. Raters in this study stated, through Content Validity testing (Table 4), that they preferred to use an assessment scale with a range of performance, thus the dichotomous *Yes* or *No* scale was changed to a Likert scale to allow the raters to determine the evidence as “*Excellent*”, “*Good*”, “*Fair*”, or “*Insufficient*”. The Likert scale offers a gradation in the assessment, which the dichotomous scale did not. The dichotomous scale permitted only two categories: competency on the item or lack of competency. The addition of multiple scale points may also lead to greater reliability amongst raters and between ratings because a competency assessment requires problem solving and application of knowledge and skills which are difficult to rate on a dichotomous *Yes* or *No* scale (Bashook, 2005).

The Likert scale also provides the assessor with more information upon which to determine the coach’s competency (Berger, 1998). Any areas of weakness that are noted by the

assessor should be explained to the learner through meaningful feedback that is directly related to areas of improvement. This will allow the learner to make strategic decisions about future professional development opportunities (Voorhees, 2001). This scale also prevents learners from memorizing a minimum number of checklist items needed to pass, instead emphasizing that the learner must develop knowledge, motivation, and personal attributes of a competent coach (van der Vleuten et al., 2010). The modification to the assessment scale in this study is supported by previous findings that a Likert scale is valuable for the development of an organizational culture of assessment (Kaslow et al., 2007), which for Gymnastics Canada is integral as this competency assessment is their first.

iv. Competency Assessment. In competency assessment, the results must be meaningful to the learner and be clearly displayed so they understand their strengths and weaknesses. The original Artistic Gymnastics Community Sport coaching assessment did not include an overall assessment of competency. To deliver a more meaningful assessment of competency, an overall competency assessment was added to the instrument based on the coaching development continuum: *Beginner*, *Competent*, *Proficient* and *Expert* (Schempp, et al., 2006). This continuum of performance shows differences between Beginner and Expert performances that reflect the desired competency for Gymnastics Canada coaches to “develop, implement and evaluate a safe, age and level-appropriate gymnastics lesson” (Gymnastics Canada, 2011).

Following data collection for content validity testing in Stage 2, a discussion took place between the researcher and the coach education director at Gymnastics Canada to verify if any items were mandatory for competency. It was pointed out to the researcher that Gymnastics Canada considered three situations in which a coach would be incompetent:

- The lesson shows that the coach is not ready to be certified.
- The lesson has issues concerning safety.
- The lesson shows harmful or unethical coaching behaviours.

(E.Bureaud, personal communication, June 20, 2013)

Based on these situations, an assessor could request that the coach complete additional training and be re-assessed. This would fill the gap identified by raters through Stage 2 content validity testing (Table 4) that no known parameters for re-assessment were included in the original instrument.

In the original instrument, evidence of coaching competency was based on the observation of performance. The addition of an overall assessment following Stage 2 provides much needed transparency for learners to know what is expected of their performance (Linn et al., 1991) and to gain awareness of the variations in performance that would cause them to be successful or unsuccessful (Bowden, 2000). More research is recommended to verify if gymnastics experts agree on the descriptions of Community Sport coaching expertise on the continuum.

It is also recommended that Gymnastics Canada publish a coach development pathway for coaches to understand the quality of their performance on the continuum of expertise. This would aid beginner coaches to identify professional development activities or practical applications needed to advance to the expert level in the context. The inclusion of competency levels in this assessment may also guide Gymnastics Canada to review their objectives of competency assessment, measure the progress of coaches more carefully, and align them to organizational strategic goals such as increasing the quality of coaching.

v. Expert Coach Participants. To determine content validity of the assessment items, the opinions of experts were surveyed. They provided feedback on areas for improvements within the instrument, including their views of the performance rating scale. Although all raters were certified by Gymnastics Canada as NCCP Learning Facilitators of coach education, their personal traits and years of experience were unknown to the researcher. Gymnastics Canada should ensure that future recruitment of experts for assessment follows the descriptions of expertise on the coaching development continuum. To do this, recruitment of raters requires selection of the best representatives from not only experience in the context, but mastery and demonstration of expert knowledge and skills. In addition, the raters must portray attitudes and values that match those of the organization (Patterson et al., 2000). They are in the position to lead the organization in coach education, training and assessment by adopting and modeling the desired culture. This requires equivalent training of raters to identify coaching competency and to mentor coaches who are below the minimum performance level (Lichtenberg, et al., 2007).

4.3 Stage 3

Three regional coaches provided their views of the instrument's applicability to coaching Community Sport Artistic Gymnastics. The feedback confirmed face validity of the modified

instrument, meaning it was credible and reflected the context in which they are currently coaching (Nevo, 1985).

Despite receiving information about the goals of competency training and assessment, none of the coaches recommended any further modifications to the instrument. This may have been due to the fact that content validity testing by the expert raters occurring first, allowing the expert group to identify the necessary changes. The coaches may have lacked expertise to provide insight beyond what the experts viewed in regards to the essentiality of the items on the instrument.

The coaches had a minimum of five years of coaching experience and certification in the NCCP Community Sport context. Given the recent transition to a CBET model for gymnastics coaches in Canada, these coaches did not experience competency training or assessment in their development. Therefore, it is unknown what level of expertise these coaches had, or what new contributions they could make to the items, the instrument, or to assessing competency.

4.4 Stage 4

Percent of Agreement overestimates agreement by not accounting for ratings that agree by chance. When ratings by chance are taken into account by AC_1 , the measured value is considerably lower than *Percent of Agreement*. This was true in the results displayed in Table 7, where the modified instrument showed two levels of agreement: *Moderate* ($AC_1 = 0.44$) and *Substantial* (*Percent of Agreement* = 0.72).

An overall (formative) assessment of competency was included in the modified instrument to identify competency on a continuum of expertise (Schempp et al., 2006). Nine of the ten raters gave the coach the minimum level for certification of *Competent*. The ability of the raters to agree more often on ratings of overall competency rather than individual item ratings is consistent with previous inter-rater reliability findings for competency assessments (Baartman et al., 2007; Bakker et al., 2011). However, only four of those nine raters agreed that the coach was *Competent* while three raters said the coach was *Proficient* and two raters viewed the coach as *Expert*. These findings indicate that raters rated the coach at the minimum level of competency, but they did not agree on the coach's level of expertise.

In competency assessment this overall rating is of greater value than the item level reliability, as the objective of a competency assessment is “to determine whether a defined level of mastery has been achieved” (Miller, 1990, p. S66). The challenge to achieve agreement of expertise ratings begins with the raters understanding the objective of the assessment and the expertise categories (Mellenbergh, 2011). The raters may have focused on rating individual items and disregarded the purpose of the ratings to be used to determine the coach’s competency.

Understanding that the instrument is designed to assess coaching competency for certification, it would be in the best interest of Gymnastics Canada to improve the inter-rater reliability of the instrument before using it as the means to award (or not award) coaching certification. At minimum, a *Substantial* level of agreement creates confidence in the conclusions based on the use of this instrument, while *Almost Perfect* agreement would guarantee reliability of the raters and the conclusions about the coach from the use of the instrument (Gwet, 2010). Improvements should be made towards a minimum of *Substantial* level of agreement when measured by the Agreement Coefficient ($AC_1 \geq 0.61$) or *Almost Perfect* agreement when measured by Percent of Agreement ($PA \geq 0.81$) for item level ratings and for ratings of the coach’s level of expertise.

The moderate inter-rater reliability reveals an inability of the raters to rate the coach’s performance on the assessment scale (Kottner et al., 2011; Murphy & Davidshofer, 2001). To achieve a higher level of agreement, the raters should be educated on the item level differences on the scale and receive practice in rating video samples and using the instrument (Brewer & Jones, 2002; Cushion et al., 2012). This would require a collection of coaching samples along the scale for each item in the instrument, from *Insufficient* to *Excellent*.

Assessors should be trained to focus only on the validated competencies and be able to provide ratings consistent with the group of assessors. A practical rater training should reduce personal bias of the raters, an inherent characteristic of expert judgments (van der Vleuten, et al., 2010). The training should include a reliability test to measure the level of agreement amongst the raters, with a pre-determined minimum level of agreement to become an official assessor (e.g. 0.81 Percent of Agreement) (Hallgreen, 2012). Raters who do not meet the minimum level of agreement should receive more training, or be removed from the group of assessors (Mellenbergh, 2011).

The more objective the raters are, the more likely they are to use evidence to rate the coach's competencies and to determine the coach's level of expertise (Hallgreen, 2012). The training should include an opportunity for raters to discuss differences in opinions on individual items, specifically the items that received at least one rating within each possible category of "*Insufficient*", "*Fair*", "*Good*", and "*Excellent*". The variance may have been due to incorrect or unclear wording, subjectivity of the item, or difficulty of the item to be observed (Traub & Rowley, 1991; Voorhees, 2001). All are challenges that could be identified during detailed rater training.

4.5 Stage 5

The intra-rater reliability demonstrates the ability of the raters to provide *Substantial* and *Almost Perfect* levels of agreement in two subsequent observations. The strong agreement between ratings by individuals shows that scoring error is minimal for the same rater on multiple occasions.

One factor that may have contributed to the strong intra-rater statistics is that, despite a week-long time lapse, raters could have remembered their own ratings or the events they viewed the first time (Traub & Rowley, 1991). There also may have been a learning effect leading to the improvement of intra-rater reliability. To alleviate these effects, the time lapse could have been lengthened.

The inter-rater reliability results in Stage 4 compared to the intra-rater reliability results in Stage 5 indicate that for the modified instrument the chance of error in ratings is more likely to occur between multiple raters than within one rater. The capacity of raters to provide repeat testing results shows clarity of the assessment items, which is benefitted by raters understanding the assessment objectives, content and procedures (Traub & Rowley, 1991; Voorhees, 2001). Efforts to enhance the instrument's reliability should focus on improving agreement between raters rather than through changes to the assessment instrument (van der Vleuten, et al., 2010). Attention should be placed on the eliminating causes of poor inter-rater reliability before using the instrument to assess gymnastics coaches for competency.

4.6 Future Directions of Coaching Competency

i. Coaching Development. Canada's National Coaching Certification Program (NCCP) is known to be valuable for improving coaching quality (Misener & Danylchuk, 2009). Unfortunately it is unknown if this value exists after the recent change from knowledge-based coach education to Competency-Based Education and Training (CBET). The Coaching Association of Canada (CAC), as the regulating body of the NCCP, has attempted to expand the depth of training they provide to meet their organizational objective to "lead training and development of competent coaches" (CAC, 2011, p.1). As the NCCP has only recently integrated CBET requirements (CAC, 2011a), the traditional understanding of coach development has been challenged by recent emphasis on what coaches do with what they know, rather than just what coaches know (Miller, 1990; Santos, 2010). CAC should investigate the values coaches place on the current NCCP in a CBET model, and if they believe their coaching quality and competency improved through the program.

In order for the NCCP's CBET program to elicit positive results, all members of the coaching community including the coaches, educators, administrators, and assessors must understand the objectives of the program. The findings in this study about the Gymnastics Canada competency assessment as a valid and moderately reliable instrument are good baselines from which to generate discussions of the value of CBET for developing quality coaches and creating positive athlete experiences.

The Community Sport context was Gymnastics Canada's first context to receive adjustments and additions to meet CBET criteria. This study endorses their competency assessment instrument for Community Sport coaches and allows them to recommend that member coaches complete the assessment to receive formal certification. The organization's commitment to a valid and reliable assessment displays their interest in fostering a culture of competency training and assessment, which is required for a successful CBET program (Kaslow et al., 2007). By establishing the validity of their assessment instrument, Gymnastics Canada can focus on improving the delivery of learning opportunities to develop coaching competency and improving the reliability of assessors to achieve the NCCP goal of enhanced coaching quality.

ii. Competency Assessment. In Stage 4, the coach's competency was rated in all four possible levels of expertise from *Incomplete* to *Expert*. Despite the Gymnastics Canada assessment instrument being valid, the reliability of this rating could be higher. The consistency of the raters may have been limited by the quality of evidence provided by the coach. In a single lesson, the coach may not demonstrate enough context-specific knowledge and skills for the raters to use to determine a level of expertise (Gibb, 1993; Hay et al., 2012). In the future, coaches who are selected as models in the standardization and reliability testing of a CBET assessment instrument should be aware, prior to conducting their lesson, of the assessment items, the rating categories, and the objective of the assessment. It is also recommended that a discussion between the coach and the assessor follow the observation (Bakker et al., 2011; Hay et al., 2012). In this study, observers did not meet the coach nor engage in dialogue to gain insight into the rationale for the coach's behaviours in the video. These discussions are also beneficial for coaches to learn to reflect on their actions, to become conscious of their behaviours, and to provide rationale for their actions (Cushion et al., 2003).

It is well established that successful coaches develop by learning through several sources: formal, non-formal, and informal (Nelson et al., 2006). Competency assessment awards certification based on the coach's ability, regardless of whether the coach acquires experience and skills formally or informally. Due to the contextual nature of sports coaching it is imperative to identify context-specific competencies prior to implementing competency training and assessment (ICCE, 2012). Kaslow et al. (2007) recommends a working group of professionals in the field come to a consensus in the definition of competencies in all stages of professional development. Although CAC has done similar working group definitions for the core competencies of coaches (CAC, 2010), and Gymnastics Canada defines their recreational context competencies through the validity stages of this study, it is crucial that different sports, or contexts of the same sport, identify sport specific competencies.

Learners say that a competency-based model for training and assessment was a positive experience when their developmental activities led to successful attainment of their goals (De Vos et al., 2011). It is recommended that organizations, such as Gymnastics Canada, frequently review their desired coaching competencies. They could then plan for future alterations to assessments to account for evolution and adaptation in both the sport and in coaching. The

competencies to be assessed should align with the instructional components of the CBET model to show that changes in the learner's behaviour have occurred as a result of the education and training (Bowden, 2000). After selecting the desired competencies, an evaluation process, such as Brewer and Jones' (2002) five-stage process that was followed in this study, can determine whether a successful assessment framework has been established (Baartman et al., 2006; Rodriguez et al., 2002). The process used to determine the validity and reliability of the assessment can be applied to assessments in any sport or context. Noting that a single sport (gymnastics) and single context (recreation) was studied here, further research of coaching competency assessments is encouraged for national coaching certification programs.

4.7 Limitations

A limitation of this study is the sample sizes of five experts and three coaches are potentially too few. With seven or more content experts for a Content Validity Index, more than one rater's opinion is needed to remove an item from the instrument, whereas a single rating from one of five experts was sufficient to remove the item in this study (Devon et al., 2007; Lynn, 1986). The possibility to use seven or more experts in the future is feasible as ten experts were recruited out of a large national pool of candidates.

Although experts are believed to have an extensive knowledge base in the sport and context, it is unclear whether this knowledge allows them to discriminate the quality of the coaching they see during an assessment. The criteria to recruit experts for inter-rater reliability testing could be performance oriented, and rather than be based upon years of experience and NCCP learning facilitator status, selection could be stronger if it was based on a demonstration of their expertise in the context including their ability to transfer their knowledge into meaningful learning opportunities for beginner coaches.

Face validity is a subjective interpretation by lay people, and provides weaker results about the ability of the instrument to measure the desired construct (DeVon et al., 2007; Nevo, 1985). It is, therefore, less robust than validity results established by experts in the field. In this study, the qualitative feedback provided by three context-specific coaches provided no new information about the instrument. It is possible that the instrument would have been scrutinized more rigorously if more coaches were surveyed. The coaches who were surveyed were certified

in NCCP prior to the implementation of a Competency-Based Education and Training (CBET) program. whereby inclusion of coaches who receive CBET training would provide feedback about an instrument in which they hold greater value for their own experience as it is required for their desired professional certification.

Another limitation of this study is the expert group's lack of understanding of competency as a holistic ability and interrelationship of an individual's knowledge, skills, and personal attributes in a given environment (Voorhees, 2001). Since human behaviour is complex, and coaching requires problem solving, there are moments in all scenarios that are unpredictable, even for expert observers (Schembri, 1995). This may have led experts to focus on singular skills and tasks while neglecting to view the whole coach. The raters may have had difficulty in observing the coach, the athletes, or the overall environment given the observation was through video and not in person. This may have led to disagreements in ratings between raters, or within raters from one observation to the next.

It is reported that observing single lessons is insufficient for observers to make a decision about the individual's competency (Sonnadara et al., 2013; Voorhees, 2001). This may have contributed to the wide variance of ratings of the coach as Beginner, Competent, Proficient, and Expert. Raters have also indicated that using multiple observations over time allows them to see improvements in performance (Bakker et al., 2011).

From a statistical perspective Percent of Agreement is easier to understand (most people can interpret a value on the 100-point scale) than the newer Agreement Coefficient that requires the use of specially designed statistical software, and the ability to interpret a value that accounts for chance of agreement. Therefore, comparisons from the inter-rater reliability in this study to others are limited to results found using the same statistical methods.

4.8 Future Research

With the growing trend for national coaching certification programs to be competency-based, more research in training and assessing coaching competency is warranted. The majority of evidence for competency-based training and assessment is rooted in research from professions of medicine, teaching and nursing. Although significant work has been done to determine

effective behaviours of sport coaches, there are few investigations of competency development and assessment.

There remains a need for evidence of how coaches develop expertise; particularly in early levels of sport as the majority of coach learning research focuses on coaches working with athletes in elite or competitive contexts (Cushion et al., 2010). Determining the barriers coaches encounter in their development of expertise is valuable for adjusting individual development pathways that lead to success. Through validating an assessment instrument in this study, the opportunity exists to repeat coach assessments year after year (or as frequently as desired) to see changes in coaching over time. This consistent feedback and focus on learning could lead to greater advancements for coaches striving for, or maintaining, expertise.

This study evaluated a competency assessment instrument for coaches in a single sport (gymnastics) and context (recreational). Given that most coach observation instruments are not sport or context specific, their use is not transferrable to all coaching behaviours, and therefore not beneficial to developing coaches in all contexts (Brewer & Jones, 2002). CAC member sports offering NCCP certification should be urged to utilize sport and context specific assessments that are both valid and reliable since coaching effectiveness is sport and context dependent (Côté & Gilbert, 2009). This suggests that each sport needs to produce a competency-based training and assessment program for certification, and to follow a similar five-stage process of establishing validity and reliability of their instruments. Doing so would allow sport organizations to align their CBET activities to their assessment and discuss coach development pathways with international partners in the united effort to advance coaching as a profession.

5. CONCLUSION

By following the previously published five-stage process to establish validity and reliability, this study is the first to evaluate a coaching competency assessment instrument in the National Coaching Certification Program (NCCP). Sport organizations that are required to use competency models must be accountable for the quality of their assessment instruments for awarding coaching certifications. This study advances the evidence for CBET and assessment for coaches by systematically confirming the validity and reliability of coaching competency assessment in the recreation sport context of Artistic Gymnastics. While the instrument's validity is confirmed, several recommendations are included to improve its inter-rater reliability.

Experts confirm the modified competency assessment instrument for Community Sport Artistic Gymnastics is relevant, practical and appropriate for the coaching context. Considering that the process in designing the original instrument is unknown, the item-level Content Validity Index confirms the relevance of forty-seven items that are essential requirements for coaches in the context. To clarify the rating process of each item, a scale of multiple points, rather than a dichotomous scale, is preferred. This is because it broadens the options for raters to select a level of performance, while increasing feedback for the coach of what constitutes unsuccessful versus successful performance. The ratings can be in one of four possible levels: *“Insufficient”*, *“Fair”*, *“Good”*, or *“Excellent”*.

This evidence guides the assessor's rating of the coach's competency as a level of expertise, which is needed to determine if the coach meets the level of performance required for accreditation. Therefore, detailed evidence is required to make a decision of the learner's competency based on their performance (Kaslow et al., 2007). Like the item-level rating scale, this overall rating is valuable for both the raters and the coach to understand the differences between beginner coaches, who are not yet competent, and those who are competent and therefore deserving of certification in the context.

To assign an overall competency rating, the raters use four stages of coaching expertise from Beginner, Competent, Proficient, to Expert. In this structure, assessors can use the evidence from the observation to guide their summative assessment of the interrelationship of the coach's knowledge, skills and personal characteristics as seen in their actions and decisions (Voorhees, 2001). The descriptions of these stages of expertise are useful to prompt action plans

for future developmental, especially when the learner recognizes that competency is only one stage of expertise and that, after certification, learning does not end (Voorhees, 2001). An overall competency rating is known to lead to higher reliability as opposed to using a checklist of tasks for observations of complex tasks (Bashook, 2005).

Moderate levels of inter-rater reliability were found for the modified instrument when a coach was observed and rated by ten raters. When five of the raters re-rated the coach, their intra-rater reliability met levels of moderate, substantial and almost perfect agreement between their two sets of ratings. These findings show that the same raters were able to use the instrument and provide consistent ratings, but that greater attention should be paid to how the raters concluded their ratings. The inconsistencies could be due to the use of videos to observe the coaching performance, the preparation of the raters, or issues with rating ambiguous or inconsistent evidence (Bakker et al., 2011).

Gymnastics Canada, along with provincial gymnastics associations and coach educators, are leaders in designing and delivering coach education. They provide the foundation of training through educational and skill-based courses and access to resources. Despite their leadership efforts, the responsibility of a coach's development ultimately rests with the individual and their desire to advance through stages of development and access resources to assist them in learning (Evers et al., 1998). The competency-based approach encourages individuals to take the time and procedures necessary to develop and recognizes that individuals learn at different rates. Because gymnastics coaches are primarily working in a club environment, they should be provided with accessible learning opportunities from formal (NCCP courses) and informal (interactions with peers) sources that support the development of essential learning outcomes for certification as determined by experts in their context (De Vos et al., 2011).

This study confirms the value in researching the validity and reliability of coaching competency assessments for certifications. Although coaching is seen as a vocation, each individual coach determines the value they place on education, training, and assessment in alignment with their vocational or professional goals. Further studies are needed on the role of assessment in the development of coaches, and how a CBET model positively influences both coaching performance and retention of coaches. As this study only analyzes coaching in the recreational context of gymnastics, other sports are encouraged to verify their assessments for

certification in a similar process. More research is also needed to show if the CBET approach to coach education is more effective than previous knowledge-based training, and whether coaches are accessing primarily formal or informal learning opportunities in the development of sport-specific coaching competency.

6. REFERENCES

- Abraham, A. and Collins, D. (1998). Examining and extending research in coach development. *Quest*, 50, 59-79.
- Abraham, A., Collins, D., & Martindale, R. (2006). The coaching schematic: Validation through expert coach consensus. *Journal of Sports Sciences*, 24(6), 549-564.
- Advanced Analytics. (2011). AgreeStat Software (Version 2011.2) [Software]. Available from: <http://www.agreestat.com>
- Baartman, L.K.J, Bastiaens, T.J., Kirschner, P.A., & van der Vleuten, C.P.M. (2006). The wheel of competency assessment: Presenting quality criteria for competency assessment programs. *Studies in Educational Evaluation*, 32, 153-170.
- Baartman, L.K.J, Bastiaens, T.J, Kirschner, P., & van der Vleuten, C.P. (2007). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. *Educational Research Review*, 2(2), 114-129.
- Baker, E.L., O'Neil, H.F., & Linn, R.L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist*, 48(2), 1210-1218.
- Bakker, M., Roelofs, E., Beijaard, D., Sanders, P., Tigelaar, D., Verloop, N. (2011). Video portfolios. The development and usefulness of a teacher assessment procedure. *Studies in Educational Evaluation*, 37, 123-133.
- Bashook, P.G. (2005). Best practices for assessing competence and performance of the behavioral health workforce. *Administration and Policy in Mental Health*, 32(5/6), 563-592.
- Becker, A. (2009). It's Not What They Do, It's How They Do It: Athlete Experiences of Great Coaching. *International Journal of Sports Science & Coaching*, 4 (1), 93-119.
- Bergmann Drewe, S. (2000). Coaches, ethics and autonomy. *Sport, Education and Society*, 5(2), 147-162.
- Berliner, D.C. (1988). *The development of expertise in pedagogy*. Paper presented at the Charles W. Hunt Memorial Lecture at the annual meeting of the American Association for Colleges of Teacher Education, New Orleans, LA.
- Blank, R.K. (2013). What research tells us: Common characteristics of professional learning that leads to student achievement. *Journal of Staff Development*, 34(1), 50-53.
- Blank, W.E. (1982). Handbook for developing competency-based training programs. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

- Booth, R. (2000). Competency based assessment – one minute wonder or here to stay? Practitioners' attitudes to assessment in 2000. *Australian Vocational Education and Training Research Association 2000 Conference*. 23-24 March.
- Bowden, J.A. (2000). Competency-based education: Neither a panacea nor a pariah. *Proceedings of Technological Education and National Development 97 Conference*, Abu Dhabi.
- Bowen, D.E., Ledford, G.E., & Nathan, B.R. (1991). Hiring for the organization, not the job. *The Executive*, 5(4), 35-51.
- Brewer, C.J., & Jones, R.J. (2002). A five-stage process for establishing contextually valid systematic observation instruments: The case of rugby union. *The Sport Psychologist*, 16, 138-159.
- Buhai, S. L. (2012). Profession: a definition. *Fordham Urban Law Journal*, 40(1), 241.
- Cassidy, T., Jones, R., & Potrac, P. (2009). Understanding sports coaching: The social, cultural and pedagogical foundations of coaching practice (2nd ed.). London: Routledge.
- Cassidy, T., & Kidman, L. (2010). Initiating a National Coaching Curriculum: A Paradigmatic Shift? *Physical Education and Sport Pedagogy*, 15 (3), 307-322.
- Coaching Association of Canada (CAC), 2005. Summary of the NCCP evaluation project: A blueprint for change.
- Coaching Association of Canada, (2010). NCCP Policies and Implementation Standards.
- Coaching Association of Canada (CAC), (2011a). 2010-2011 Annual Report.
- Coaching Association of Canada (CAC), (2011b). NCCP Evaluation Toolkit.
- Coaching Association of Canada (CAC), n.d. New NCCP Generic Overview.
- Comeau, G.S. (2012). The evolution of Canadian sport policy. *International Journal of Sport Policy and Politics*, 5(1), 73-93.
- Côté, J. and Sedgwick, W. (2003). Effective behaviors of expert rowing coaches: A qualitative investigation of Canadian athletes and coaches. *International Sports Journal*, 7(1), 62-77.
- Côté, J., & Gilbert W. (2009). An Integrative Definition of Coaching Effectiveness and Expertise *International Journal of Sports Science & Coaching*, 4(3), 307-323.
- Cushion, C., Armour, K., & Jones, R. (2003). Coach education and continuing professional development: Experience and learning to coach. *Quest*, 55(3), 215-230.

- Cushion, C., Harvey, S., Muir, B., Nelson, L. (2012). Developing the coach analysis and intervention system (CAIS): Establishing validity and reliability of a computerised systematic observation instrument. *Journal of Sports Sciences*, 30(2), 203-218.
- Cushion, C., Nelson, L., Armour, K., Lyle, J., Jones, R., Sandford, R., O'Callaghan, C. (2010). Coach Learning and Development: A Review of Literature. SportsCoach UK.
- De Marco, G.M., & McCullick, B.A. (1997). Developing coach expertise: Learning from the legends. *Journal of Physical Education, Recreation and Dance*, 68(3), 37-41.
- Demers, G. Woodburn, A.J., Savard, C. (2006). The development of an undergraduate competency-based coach education program. *The Sport Psychologist*, 20, 162-173.
- Devon, H.A., Block, M.E., Moyle-Wright, P., Ernst, D.M., Hayden, S.J., Lazzara, D.J., Savoy, S.M., Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship*, 39(2), 155-164.
- De Vos, A., De Haw, S., & Van der Heaven, B. (2011). Competency development and career success: The mediating role of employability. *Journal of Vocational Behavior*, 79, 438-447.
- Duffy, P., Hartley, H., Bales, J., Crespo, M., Dick, F., Vardhan, D., Nordmann, L., & Curado, J. (2011). Sport coaching as a 'profession': challenges and future directions. *International Journal of Coaching Science*, 5(2), 93-123.
- Dubois, D.D., & Rothwell, W.J. (2004). Competency-Bases Human Resource Management. Davis-Black Publishing: Palo Alto, California.
- Ericsson, K.A. & Lehmann, A.C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, 47, 273-305.
- Erickson, K., Bruner, M., MacDonald, D., & Côté J. (2008). Gaining insight into actual and preferred sources of coaching knowledge. *International Journal of Sports Science & Coaching*, 3(4), 527-538.
- Erickson, K., Côté, J., & Fraser-Thomas, J. (2007). Sport experiences, milestones, and educational activities associated with high-performance coaches' development. *The Sport Psychologist*, 21, 302-316.
- Evers F., Rush, J. & Berdrow, I. (1998). The Bases of Competence: Skills for Lifelong Learning and Employability. Jossey-Bass Publishers: San Francisco.
- Fan, X., Chen, M. (2000). Published Studies of Interrater Reliability Often Overestimate Reliability: Computing the Correct Coefficient. *Educational and Psychological Measurement*, 60, 532-542.

- Ford, P., Coughlan, E., & Williams, M. (2009). The expert-performance approach as a framework for understanding and enhancing coaching performance, expertise and learning. *International Journal of Sports Science and Coaching*, 4(3), 451-463.
- Furr, R.M., Bacharach, V.R. (2008). *Psychometrics: An introduction*. Thousand Oaks, CA: Sage Publications Inc.
- Gibb, J. (1993). Trainers/assessors and learners: The human face to the new CBVET system. *Testing Times*, 57-65.
- Gilbert, W.D. (2006). Introduction to Special Issue: Coach Education. *The Sport Psychologist*, 20, 123-125.
- Gilbert, W., Côté, J., & Mallett, C. (2006). Developmental paths and activities of successful sport coaches. *International Journal of Sports Science and Coaching*, 1(1), 69-76.
- Gilbert, W., Dubina, N, Emmett, M. (2012). Exploring the potential of assessment efficacy in sports coaching. *International Journal of Sports Science and Coaching*, 7(2), 211-214.
- Gilbert, W., & Rangeon, S. (2011). Current directions in coaching research. *Revista IberoAmericana de Psicología del Deporte (RIPED)*, 6(2), 217-235
- Gilbert, W., & Trudel, P., (1999). An evaluation strategy for coach education programs. *Journal of Sport Behavior*, 22(2), 234-250.
- Gilbert, W., & Trudel, P., (2001). Learning to coach through experience: Reflection in model youth sport coaches. *Journal of Teaching in Physical Education*, 21, 16-34.
- Gilbert, W., & Trudel, P., (2004). Analysis of Coaching Science Research Published From 1970-2001. *Research Quarterly for Exercise and Sport*, 75(4), 388 – 399.
- Green, M. (2007). Olympic glory or grassroots development?: Sport policy priorities in Australia, Canada and the United Kingdom, 1960-2006. *The International Journal of the History of Sport*, 24(7), 921-953.
- Gwet, K. (2002a). *Kappa statistic is not satisfactory for assessing the extent of agreement between raters*. (No. Series 1). Statistical Methods For Inter-Rater Reliability Assessment.
- Gwet, K. (2002b). *Inter-rater reliability: Dependency on trait prevalence and marginal homogeneity*. (No. Series 2). Statistical Methods For Inter-Rater Reliability Assessment.
- Gwet, K. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, 61, 29-48.

- Gwet, K. (2010). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters* (2nd Edition ed.). Gaithersburg, MD: Advanced Analytics.
- Gymnastics Canada, (2011). *Gymnastics Canada Evaluation - Draft Version 7*.
- Hager P., & Gonczi, A. (1996). What is competence? *Medical Teacher*, 18(1), 15-18.
- Hallgreen, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23-34.
- Hambrick, D.Z., Oswald, F.L., Altmann, E.M., Meinz, E.J., Gobet, F., Campitelli, G. (2013). Deliberate practice: Is that all it takes to become an expert? *Intelligence*. Advance online publication. doi: 10.1016/j.intell.2013.04.001.
- Hay, P., Dickens, S., Crudgington, B., Engstrom, C. (2012). Exploring the potential of assessment efficacy in sports coaching. *International Journal of Sports Science & Coaching*, 7(2), 187-198.
- Hayes, A.F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77-89.
- Heere, B., Walker, M., Gibson, H., Thapa, B., Geldenhuys, S., Coetzee, W. (2013). The power of sport to unite a nation: the social value of the 2010 FIFA World Cup in South Africa. *European Sport Management Quarterly*, 13(4), 450-471.
- Hodges, B., Regehr, G., McNaughton, N., Tiberius, R., Hanson, M. (1999). OSCE checklists do not capture increasing levels of expertise. *Academic Medicine*, 74, 1129-1134.
- Hoffman, T. (1999). The meanings of competency. *Journal of European Industrial Training*, 23(6), 275-285.
- Jones, R. (2007). Coaching redefined: an everyday pedagogical endeavor. *Sport, Education and Society*, 12(2), p. 159-173.
- The International Council of Coach Education (ICCE). (2010). Building the coaching community across the globe: A strategy for the International Council for Coach Education for the period 2010-2015. Retrieved from: http://www.icce.ws/position_papers/index.htm.
- Karelitz, T.M. (2008). How binary skills obscure the transition from non-mastery to mastery, measurement. *Interdisciplinary Research and Perspectives*, 6 (4), 268-272.
- Kaslow, N., Rubin, N., Bebeau, M., Leigh, I., Lichtenberg, J., Nelson, P., Portnoy, S., Smith, I. (2007). Guiding principles and recommendations for the assessment of competence. *Professional Psychology: Research and Practice*, 38(5), 441-451.

- Klein, A.L. (1996). Validity and reliability for competency-based systems: Reducing litigation risks. *Compensation and Benefits Review*, 28(4), 31-37.
- Kottner, J., Audige, L., Brorson, S., Donner, A., Gajewski, B., Hrobjartsson, A., Roberts, C., Shoukri, M., Streiner, D. (2011). Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *International Journal of Nursing Studies*, 48, 661–671.
- Kundel, H.L., & Polansky, M. (2003). Measurement of observer agreement. *Radiology*, 228(2), 303-308.
- Landis, J. R., & Koch, G. G. (1977). The measurmenet of observer agreement for categorical data. *Biometrics*, 33, 159.
- Layton, R. (1994). *Can the coach coach? The role of assessment in competency based courses*. Proceedings from the National Coaching Conference, Canberra. 97-101.
- Lemyre, F., Trudel, P. Durand-Bush, N. (2007). How youth-sport coaches learn to coach. *The Sport Psychologist*, 21, 191-209.
- Leung, W. C. (2002). Competency based medical training: Review. *British Medical Journal*, 325(7366).
- Lichtenberg, J.W., Bebeau, M.J., Nelson, P.D., Smith, I.L., Portnoy, S.M., Leigh, I.W., Rubin, N.J., Kaslow, N.J. (2007). Challenges to the Assessment of Competency and Competencies. *Professional Psychology: Research and Practice*, 38(5), 474-478.
- Ling, P. (1999). Assessing competence. *HERDSA Annual International Conference*, Melbourne.
- Linn, R.L., Baker, E.L., Dunbar, S.B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher*, 20(8), 15-21.
- Lucia, A.D., Lepsinger, R. (1999). *The Art and Science of Competency Models*. San Francisco: Jossey-Bass / Pfeiffer.
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382-386.
- Lyle, J., & Cushion, C. (2010). (Eds.). *Sports Coaching: Professionalisation & Practice*. Churchill Livingstone, Elsevier.
- Mallett, C.J., Trudel, P., Lyle, J., & Rynne, S.B. (2009). Formal vs. informal coach education. *International Journal of Sports Science and Coaching*, 4(3), 325-334.
- Mansfield, R.S. (1996). Building competency models: Approaches for HR professionals. *Human Resource Management*, 35(1), 7-8.

- McClelland, D. (1973). Testing for Competence Rather Than for Intelligence. *American Psychologist*, 28 (1), 1-14.
- Mellenbergh, G. J. (2011). *A conceptual introduction to psychometrics*. The Hague, Netherlands: Eleven International Publishing.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18, 5-11.
- Miller, G.E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine*, 65(9), S63-S67.
- Misener, K.E., & Danylchuk, K.E. (2009). Coaches' perceptions of Canada's national coaching certification program (NCCP): Awareness and value. *International Journal of Sports Science and Coaching*, 4(2), 233-243.
- MORI. (2004). *Sports Coaching in the UK*. Leeds: SportscoachUK.
- Moskal, B.M., & Leydens, J.A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, 7(10).
- Moss, P.A. (1994). Can there be validity without reliability? *Educational Researcher*, 23(2), 5-12.
- Murphy, K.R.; Davidshofer, C.O. (2001). *Psychological testing: principles and applications*. (5th Ed.). Upper Saddle River, NJ. Prentice Hall.
- Myers, N.D., Feltz, D.L., Maier, K.S., Wolfe, E.W., Reckase, M.D. (2006). Athletes' evaluations of their head coach's coaching competency. *Research Quarterly for Exercise and Sport*, 77(1), 111-121.
- Nash, C., & Sproule, J. (2009). Career development of expert coaches. *International Journal of Sports Science and Coaching*, 1(4), 121- 138.
- Nelson, L., & Cushion, C. (2006). Reflection in coach education: The case of the national governing body coaching certificate. *The Sport Psychologist*, 20, 174-183.
- Nelson, L., Cushion, C. & Potrac, P. (2006). Formal, nonformal and informal coach learning: A holistic conceptualization. *International Journal of Sports Science and Coaching*, 1(3), 247-259.
- Nelson, L., Cushion, C. & Potrac, P. (2013). Enhancing the provision of coach education: The recommendations of UK coaching practitioners. *Physical Education and Sport Pedagogy*, 18(2), 204-218.

- Nevo, B. (1985). Face validity revisited. *Journal of Educational Measurement*, 22(4), 287-293.
- Nicholson, M., Hoye, R., & Houlihan, B. (2011). *Participation in sport: International policy perspectives*. New York, NY: Routledge.
- Patterson, F., Ferguson, E., Lane, P., Farrell, K., Martlew, J., & Wells, A. A competency model for general practice: implications for selection, training, and development. *The British Journal of General Practice* 50(542), 188-193.
- Polit, D. & Beck C. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29, 489-497.
- Roberts, S., & Fairclough, S. (2012). A five-stage process for the development and validation of a systematic observation instrument: The system for observing the teaching of games in physical education (SOTG-PE). *European Physical Education*, 18(1), 97-113.
- Rodriguez, D., Patel, R., Bright, A., Gregory, D., Gowing, M.K. (2002). Developing competency models to promote integrated human resource practices. *Human Resource Management*, 41(3), 309-324.
- Santos, S., Mesquita, I., Graça, A., Rosado, A. (2010). Coaches' perceptions of competence and acknowledgement of training needs related to professional competences. *Journal of Sports Science and Medicine*, 9, 62-70.
- Schembri, G. (1995). Coach education: The next big step. *The Coaching Director*, 8, 52- 58.
- Schempp, P., McCullick, B., & Mason, L. (2006). The development of expert coaching. In R.L. Jones, *The sports coach as educator: Re-conceptualising sports coaching*. London; New York: Routledge.
- Schempp, P., Tan, S., Manross, D., Fincher, M. (1998). Differences in novice and competent teachers' knowledge. *Teachers and Teaching: theory and practice*. 4(1), 9-20.
- Scosche. (2012). FreedomMIC for Apple iOS (Version 1.4) [Mobile application software]. Retrieved from itunes.apple.com/ca
- Sonnadara, R.R., Mui, C., McQueen, S., Mironova, P., Nousiainen, M., Safir, O., Kraemer, W., Ferguson, P., Alman, B., & Reznick, R. (2013). Reflections on competency-based education and training for surgical residents. *Journal of Surgical Education*, 71(1), 151-158.
- Sport Canada. (2012). *Canadian Sport Policy*. Retrieved from: www.pch.gc.ca/eng/1358351890624/1358352054107
- Taylor, B. and Garratt, D. (2010a). The professionalization of sports coaching: relations of power, resistance and compliance. *Sport, education and society*, 15(1), 121-139.

- Taylor, B. and Garratt, D. (2010b). The professionalization of sports coaching: definitions, challenges and critique, in: J. Lyle and C. Cushion. (Eds.) *Sports Coaching, Professionalisation and Practice*. Edinburgh: Churchill Livingstone Elsevier.
- Trudel, P., Gilbert, W., & Werthner, P. (2010). Coach education effectiveness. In J. Lyle & C. Cushion, *Sports Coaching: Professionalisation and Practice*. Churchill Livingstone, Elsevier.
- Thomas, S.D., Hathaway, D.K., Arheart, K.L. (1992) Face validity. *Western Journal of Nursing Research*, 14(1), 109-12.
- Thomson, P., Saunders, J., & Foyster, J. (2001). *Improving the validity of competency-based assessment*. Leabrook, Australia: National Centre for Vocational Education Research.
- Tracy, K., Adler, L.A., Rotrosen, J., Edson, R. (1997). Interrater reliability issues in multicenter trials. Part I: Theoretical concepts and operational procedures used in Department of Veterans Affairs Cooperative Study #394. *Psychopharmacology Bulletin*, 33(1), 53-57.
- Traub, R.E., & Rowley, G.L. (1991). Understanding reliability. *Educational Measurement: Issues and Practice*, 10(1), 37-45.
- Turner, D., Nelson, L., & Potrac, P. (2012). The journey is the destination: Reconsidering the expert sports coach. *Quest*, 64(4), 313–325.
- U.S. Department of Education, National Center for Education Statistics. (2001). *Defining and Assessing Learning: Exploring Competency-Based Initiatives*. Retrieved from <http://nces.ed.gov/pubs2002/2002159.pdf>
- Vallée, C.N., & Bloom, G.A. (2005). Building a successful university program: Key and common elements of expert coaches, *Journal of Applied Sport Psychology*, 17(3), 179-196.
- van der Mars, H. (1989). Observer Reliability: Issues and Procedures. In P.Darst, D.Zakrasjek, V.Mancini, *Analyzing Physical Education and Sport Instruction* (pp. 53-78). Champaign, IL: Human Kinetics.
- van der Vleuten, C.P.M., Schuwirth, L.W.T., Scheele, F., Driessen, E.W., Hodges, B. (2010). The assessment of professional competence: building blocks for theory development. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 24(6), 703-719.
- Voorhees, R.A. (2001). Competency-based learning models: A necessary future. *New Directions for Institutional Research*, 110, 5-13.
- Warrens, M.J. (2013). Cohen's quadratically weighted kappa is higher than linearly weighted kappa for tridiagonal agreement tables. *Statistical Methodology*, 9, 440-444.

- Waltz, C.F., Strickland, O.L., Lenz, E.R. (2005). *Measurement In Nursing and Health Research* (3rd ed.). New York: Springer.
- Werthner P., & Trudel, P. (2006). A new theoretical perspective for understanding how coaches learn to coach. *The Sport Psychologist*, 20, 198-212.
- Wiman, M., Salmoni, A.W., & Hall, C.R. (2010). An examination of the definition and development of expert coaching. *International Journal of Coaching Science*, 4(2), 37-60.
- Woodman, L. The development of coach education in Australia. *The Journal of the Australian Society for Sports History*, 5(2), 204-223.
- Wongpakaran, N., Wongpakaran, T., Wedding, D., & Gwet, K. (2013). A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Medical Research Methodology*, 13, 61.

APPENDIX A



January 14, 2013

Dear Gymnastics Learning Facilitator,

As you may be aware, the National Coaching Certification Program evaluation of competency for gymnastics community coaches (Gymnastics Foundations) is still in draft stages. We are conducting a research project at the College of Kinesiology in Saskatoon to finalize this evaluation as valid (accurate) and reliable (consistent) for its use in accrediting coaches across Canada.

Based on your certification as a Learning Facilitator (LF), you are recognized as an expert in the Community Sport Stream. For this project to be successful, ten facilitators are required to assist in establishing validity and reliability of the coach competency evaluation criteria in a five-stage process. The process will require between 1.5 hours to 9 hours of your time spanning over several weeks.

The initial stages require five LFs to receive training to observe coaching videos and code the behaviours in the videos to the competency evaluation criteria. Training will follow with a validity survey. Ten LFs will participate in inter-observer reliability testing, and five of these ten will be randomly chosen to examine a test-retest reliability measure. Email will be the primary form of communication and online video conferencing will be used to view the video clips, allowing you to participate with a personal computer and an internet connection.

By participating in this research, you will make an immediate contribution to a validation process on how Gymnastics Canada assesses coaching competency. You will be a significant contributor by providing your knowledge and experience of coaching community sport gymnastics. This contribution will establish a process that will be in place for years to come and allow entry level coaches to be evaluated on their coaching competency based on a statistically reliable and valid set of criteria.

Please read the attached Consent Form. If you agree to participate, a signed copy can be returned to the email or mail contact below.

If you are interested in learning more about this study, or if you have any questions regarding the procedure, please contact me for more details.

Sincerely,

Alynn Brischuk
MSc Candidate, College of Kinesiology
University of Saskatchewan
PAC, 87 Campus Drive
Saskatoon SK S7N 5B2
alynn.brischuk@usask.ca
P: 306-966-1123
F: 306-966-6464

Dr. Keith Russell
College of Kinesiology
University of Saskatchewan
PAC, 87 Campus Drive
Saskatoon SK S7N 5B2
keith.russell@usask.ca
P: 306-966-1067
F: 306-966-6464

CONSENT FORM – EXPERT OBSERVERS

You are invited to participate in a research project entitled *Establishing a Valid and Reliable Competency Evaluation for Gymnastics Coach Certification*. Please read this form carefully, and feel free to ask questions you might have.

Researcher: Alynn Brischuk Graduate Student, College of Kinesiology
306-966-1123, alynn.brischuk@usask.ca

Supervisor: Dr. Keith Russell, College of Kinesiology
306-966-1067, keith.russell@usask.ca

Purpose:

The Coaches Association of Canada (CAC) is responsible for a national strategy to educate and evaluate coaches based on competencies. Coaches who are seeking certification are required to demonstrate these competencies as applications of knowledge and skills in sport specific environments.

The purpose of this study is to establish a reliable and valid competency evaluation for gymnastics community sport coaches. Through a five stage process, this project will confirm that the behaviour evidences on the evaluation instrument are accurate descriptions of coaching competency. In several stages, video clips of active coaches will be viewed to ensure consistent results are found among multiple expert observers.

Procedure:

The first stage will be to train expert observers (Learning Facilitators) on the list of competencies and have them practice by viewing and coding video clips. In the second stage a Content Validity Index will be completed independently by each expert observer. Later stages will examine inter and intra observer reliability among a select number of expert observers.

The researchers will videotape community sport coaches delivering recreational lessons of no longer than 60 minutes. These community coaches will provide detailed program descriptions and lesson plans. Video clips from the lessons will be viewed by expert observers and evaluated using evidences listed in the evaluation instrument. Data will be analyzed for reliability between all of the expert observers and for each individual expert observer. Community sport coaches will remain anonymous and no feedback or assessment of the coach's competency will be evaluated in this process, rather the focus of the activity is to ensure accuracy and consistency of the expert observers to view the same coaching behaviours.

Expert observations will take place online through web-based meetings, where the researcher will initiate and host meetings using Cisco Webex software. As a meeting host, the researcher can select

the video clips to be shown and can randomize them appropriately. Meetings can be held with an expert observer anywhere they are located provided they have an internet connection.

Potential Benefits:

There is no guarantee that you will personally benefit from your involvement. By participating in this research you will have an immediate effect on the implementation of a National Coaching Certification Program evaluation through analysis of its consistency and accuracy. You will provide a significant input through your knowledge and experience of coaching gymnastics at a community sport level.

We expect this research to strengthen organizational efficacy of Gymnastics Canada by guaranteeing entry level coaches a valid status of competency. This process will establish protocol to be repeated in the future for advanced coaching certifications. In addition, this process will positively support the policy change to competency based training, and become a basis for future research in coaching competency.

Potential Risks:

As an observer, you will be asked to provide input on the evidences in the observation instrument. Therefore, your participation in this study will require you to be open and candid about your knowledge of coaching. The coding data that you provide will not be shared with anyone other than the researchers.

Confidentiality:

Expert observers will remain anonymous to each other and to the coaches who participate in the video recordings. All observational data sampled will be labeled by a reference of "Observer *n*" and not by name. Consent Forms will be stored separately from the observation assessment data, so that it will not be possible to associate a name with an observation. All communication will be sent directly to the research team and will not be distributed or copied to any other participating observers.

Coaches will not be identified; rather video clips will be sorted based on the lesson and evidences. The coding and identification will only be known to the researcher and supervisor to accurately run observer reliability measurement. There is potential, as an expert in the sport, that you may recognize coaches or gymnasiums shown in the videos. These coaches are ensured that their anonymity is protected so we ask that confidentiality of the research you are involved in be held with highest integrity.

Data from this study will be used for a thesis and may be published or presented at research conferences, including to Gymnastics Canada, and in all cases your identity will remain confidential.

Storage of Data:

The coaching video files will be deleted off the camera and transferred to a computer for storage. Video files will be edited into smaller clips and stored until used for analysis. Observations will take place through a secure web-based conferencing tool and at that time the researchers will display the videos. Files are not stored online or viewable by anyone other than who is intended to see them. All coding data and surveys will be stored in a password protected account and will be remain property of the College of Kinesiology for a minimum of five years as per U of S research guidelines.

Right to Withdraw:

Your participation is voluntary, and you can choose to provide input only at a time when you are comfortable doing so. The information that is shared will be held in strict confidence and discussed only with the research team. You may withdraw from the research project for any reason, at any time, without penalty of any sort. Your decision to withdraw will not impact your coaching certification status, employment status or future recruitment in similar studies.

Participants who withdraw will have no further contact with the research team and their data collected up until the time of withdrawal can be destroyed, if requested, prior to being observed for statistical analysis. Video files that have already been used for inter/intra-observer reliability analysis will be retained.

Questions:

If you have any questions concerning the research project, please feel free to ask at any point by contacting the researcher at the information provided.

The proposed research project was reviewed and exempt from ethical approval based on the Tri-Council Research Policy of Program Evaluation.

Consent to Participate:

- I have read and understood the description provided.
- I have had an opportunity to ask questions and my questions have been answered.
- I consent to participate in the research project, understanding that I may withdraw my consent at any time.
- A copy of this Consent Form has been given to me for my records.

(Name of Participant)

(Date)

(Signature of Participant)

(Signature of Researcher)

APPENDIX B



Alynn Brischuk
College of Kinesiology
Physical Activity Complex
87 Campus Drive
Saskatoon SK, S7N 5B2

December 13, 2012

To: Gymnastics Saskatchewan Level 1 and Level 2 Coaches

Dear Coaches,

As you may be aware, the National Coaching Certification Program evaluation of competency for gymnastics community coaches (Gymnastics Foundations) is still in draft stages. We are conducting a research project at the College of Kinesiology in Saskatoon to finalize this evaluation as valid (accurate) and reliable (consistent) for its use in accrediting coaches across Canada.

For this project to be successful, four coaches who have a minimum of 5 years of coaching experience at the recreational level are needed. Coaches who volunteer will be video recorded at their own gymnasium with a group of participants (recreational/non-competitive only) during a lesson of 45-60 minutes of duration.

The evaluation criteria will be validated in a five step process by nationally recruited Learning Evaluators (who will be Expert Observers). Your lesson will be coded and data will be compiled by testing reliability between all of the expert observers and for each individual expert observer to confirm the appropriateness of the observation instrument. There will also be an opportunity for you to provide input and feedback about the evaluation criteria through a survey.

By participating in this research, you will make an immediate contribution to a validation process on how Gymnastics Canada assesses coaching competency. You will be a significant contributor by providing your knowledge and experience of coaching recreational gymnastics. This contribution will establish a process that will be in place for years to come and allow entry level coaches to be evaluated on their coaching competency based on a statistically reliable and valid set of criteria.

Please read the attached Consent Form. By agreeing to participate, please return a signed copy to the contact information below and you will be contacted to schedule the video recording. Please note that the athletes within the lesson will also need to consent to participate.

If you are interested in learning more about this study, or if you have any questions regarding the procedure, please contact me for more details:

Alynn Brischuk
MSc Candidate, College of Kinesiology
PAC, 87 Campus Drive
Saskatoon SK S7N 5B2
alynn.brischuk@usask.ca
Ph: 306-966-1123
F: 306-966-6464

CONSENT FORM – PARTICIPATING COACHES

You are invited to participate in a research project entitled *Establishing a Valid and Reliable Competency Evaluation for Gymnastics Coach Certification*. Please read this form carefully, and feel free to ask any questions you may have.

Researcher: Alynn Brischuk, Graduate Student, College of Kinesiology
306-966-1123, alynn.brischuk@usask.ca

Supervisor: Dr. Keith Russell, College of Kinesiology
306-966-1167, keith.russell@usask.ca

Purpose:

The Coaches Association of Canada is responsible for a national strategy to educate and evaluate coaches based on competencies. Coaches who are seeking certification are required to demonstrate these competencies as applications of knowledge and skills in sport specific environments.

The purpose of this study is to establish a reliable and valid competency evaluation for gymnastics community sport coaches. Through a five stage process, this project will confirm that the behaviour evidences on the evaluation instrument are accurate descriptions of coaching competency. In several stages, video clips of active coaches will be viewed to ensure consistent results are found among multiple observers.

Procedure:

The researchers will videotape community sport coaches delivering recreational lessons of no longer than 60 minutes. These community coaches will provide detailed program descriptions and lesson plans. Video clips from the lessons will be viewed by expert observers and evaluated using evidences listed in the evaluation instrument. Data will be analyzed for reliability between all of the expert observers and for each individual expert observer. Community sport coaches will remain anonymous and no feedback or assessment of the coach's competency will be evaluated in this process, rather the focus of the activity is to ensure accuracy and consistency of the expert observers to view the same coaching behaviours.

Participating community sport coaches will be requested to take a maximum of one hour to answer a questionnaire regarding the face validity of the evaluation criteria to determine whether they reflect the community sport coaching.

Potential Benefits:

There is no guarantee that you will personally benefit from your involvement. By participating in this research you will have an immediate effect on the implementation of a National Coaching Certification Program evaluation through analysis of its consistency and accuracy. You will provide a significant input through your knowledge and experience of coaching gymnastics at a community sport level.

We expect this research to strengthen organizational efficacy of Gymnastics Canada by guaranteeing entry level coaches a valid status of competency. This process will establish protocol to be repeated in

the future for advanced coaching certifications (Level 2 and higher). In addition, this process will positively support the policy change to competency based training, and become a basis for future research in coaching competency.

Potential Risks:

Your participation in this study will require you to be open and candid about your coaching knowledge, skills and behaviours. This will also require you to share your coaching methods, ideas and best practices.

Trained expert observers will be viewing your coaching. Individual styles, personalities and technical knowledge will not be evaluated or measured; however, the experts may identify areas that you are or are not competent in based on the research required of them to code your behaviours to competency evidences. The coding categories will not be shared with anyone other than the researcher of this project.

You may have received training or professional development delivered by any of the expert population therefore a risk may exist that the expert observers recognize you even though your work is intended to be anonymous. There is no expectation that results can be impacted by this risk of recognition or from bias due to the study design and the statistical analysis being corrected for chance.

The standard risks involved as a sport coach are present.

Videotaping of your lesson may lead to the researcher seeing unsafe or unethical coaching behaviours. A professional level of coaching is expected. If during video collection the researcher observes unprofessional actions, they will discontinue video recording, delete the video collected until that point, and inform you in person at the end of the session that the video will not be used in the study with rationale for the decision.

Confidentiality:

Community sport coaches will not be identified; rather video clips will be sorted based on the lesson and evidences. The coding and identification will only be known to the researcher and supervisor to accurately run observer reliability measurements. Video data will be accessible only to the researcher and the research supervisor. Consent forms and video files will be stored separately so that it will not be possible to associate a name to a video.

The expert observers who will be observing and validating the instrument tool will be recruited from a pool of national Learning Facilitators. Your anonymity is a priority, yet given the small gymnastics coaching community, there may be a risk that an observer will recognize you and/or your gymnasium from your video. In addition, to keep the essence of the coaching environment, video recordings may collect incidences where your athletes refer to you by your first name. The video may be edited to remove this part of the audio if the behaviour evidence is not compromised.

All qualitative information collected from questionnaires will remain anonymous.

Data from this study will be used for a thesis and may be published or presented at research conferences, including to Gymnastics Canada, and in all cases your identity will remain confidential.

Storage of Data:

The coaching video files will be deleted off of the collection device and transferred to a computer for storage. Video files will be edited into smaller clips and stored until used for analysis. Observations will take place through a secure web-based conferencing tool and at that time the researchers will display

the videos. Files are not stored online or viewable by anyone other than who is intended to see them. All data will be stored in a password protected account and will remain property of the College of Kinesiology for a minimum of five years as per U of S research guidelines.

Right to Withdraw:

Your participation in this study is voluntary and you can withdraw at any time. You may withdraw for any reason, at any time, without penalty of any sort. Your decision to withdraw will not impact your coaching certification status, employment status or future recruitment in similar studies.

Participants who withdraw will have no further contact with the research team and their data collected up until the time of withdrawal can be destroyed, if requested, prior to being observed for statistical analysis. Video files that have already been used for inter/intra-observer reliability analysis will be retained.

Questions:

If you have any questions concerning the research project, please feel free to ask at any point by contacting the researcher at the information provided.

The proposed research project was reviewed and exempt from ethical approval based on the Tri-Council Research Policy of Program Evaluation.

Consent to Participate:

- I have read and understood the description provided.
- I have had an opportunity to ask questions and my questions have been answered.
- I consent to participate in the research project, understanding that I may withdraw my consent at any time.
- A copy of this Consent Form has been given to me for my records.

Participant to provide initials:

- Videos may be taken of me for: Analysis _____ Dissemination* _____

**Even if no names are used, you may be recognizable if visual images are shown as part of the results.*

(Name of Participant)

(Date)

(Signature of Participant)

(Signature of Researcher)

APPENDIX C



College of Kinesiology

NCCP COACHING OBSERVATION STUDY

CONSENT FORM – PARTICIPATING ATHLETES

You are invited to participate in a research project entitled *Establishing a Valid and Reliable Competency Evaluation for Gymnastics Coach Certification*. Please read this form carefully, and feel free to ask questions you might have.

Researcher: Alynn Brischuk, Graduate Student, College of Kinesiology
306-966-1123, alynn.brischuk@usask.ca

Supervisor: Dr. Keith Russell, College of Kinesiology
306-966-1167, keith.russell@usask.ca

Purpose:

Coaches who are seeking certification are required to demonstrate competency in sport specific environments.

The purpose of this study is to establish a reliable and valid competency evaluation for gymnastics community sport coaches.

Procedure:

The researchers will record coaches delivering recreational lessons of no longer than 60 minutes. Video clips from the lessons will be viewed by expert observers and coded. Data will be analyzed for reliability between all of experts.

If you agree to participate in the lesson, your coach will be the centre of the video, but your actions will be seen.

Potential Benefits:

There is no guarantee that you will personally benefit from your involvement. By participating in this research you will assist in the implementation of a National Coaching Certification Program coach evaluation system.

Potential Risks:

Your participation in this study will require you to be active in a gymnastics lesson, meaning the standard risks involved in active sport participation are present.

Confidentiality:

No personal information or contact information will be collected, and no assessment the athletes will occur. While being videotaped, the coach may give instructions on a first-name basis that could be recorded in the video. Expert observers will be selected from a pool of national Learning Facilitators; however, they will only be shown the videos (not receive copies), and videos will not be viewable by anyone but the research team and the observers.

If you agree to participate, the consent forms will be stored separately from the videos and observation data. Each observer watching the videos will be asked to respect the confidentiality of the coach and participants in the videos by not disclosing any of the contents of the videos outside of the group.

Data from this study will be used for a thesis and may be published or presented at research conferences, including to Gymnastics Canada, and in all cases your identity will remain confidential.

Storage of Data:

The coaching video files will be deleted off the camera and transferred to a computer for storage. Video files will be edited into smaller clips and stored until used for analysis. Observations will take place through a secure web-based conferencing tool and at that time the researchers will display the videos. Files are not stored online or viewable by anyone other than who is intended to see them. All data will be stored in a password protected account and will remain property of the College of Kinesiology for a minimum of five years based on U of S research guidelines.

Right to Withdraw:

Your participation in this study is voluntary, and you can withdraw at any time. Video files that have already been used for inter/intra-observer reliability analysis will be retained.

Questions:

If you have any questions concerning the research project, please feel free to ask at any point by contacting the researcher at the information provided.

The proposed research project was reviewed and exempt from ethical approval based on the Tri-Council Research Policy of Program Evaluation.

Consent to Participate:

- I have read and understood the description provided.
- I have had an opportunity to ask questions and my questions have been answered.
- I consent to participate in the research project, understanding that I may withdraw my consent at any time.
- A copy of this Consent Form has been given to me for my records.

Participant or parent/guardian to provide initials:

- Videos may be taken of me [my child] for: Analysis _____ Dissemination* _____

**Even if no names are used, you [or your child] may be recognizable if visual images are shown as part of the results.*

(Name of Participant)

(Signature of Parent or Guardian)

(Date)

(Signature of Researcher)

APPENDIX D

April 18, 2013

Dear _____,


Thank you for your commitment to participate in our study *“Establishing a Valid and Reliable Competency Assessment for Gymnastics Coach Certification”*.

You have been randomly selected from the expert volunteer group to do Stages 1 and 2 of the research process, each requiring approximately one hour of your time. A total of 6 experts have been randomly asked to complete these stages. We thank you for your interest to participate! If you have any questions following the details in this document on how to proceed with the first stage, please do not hesitate to contact us.

Stage 1 is designed for “Rater Training”. Each rater will independently view and rate the behaviours they see in the video to the coach assessment form. The assessment form to be used in training is the current draft version of the Gymnastics Foundations Evaluation (v.7, 2011).

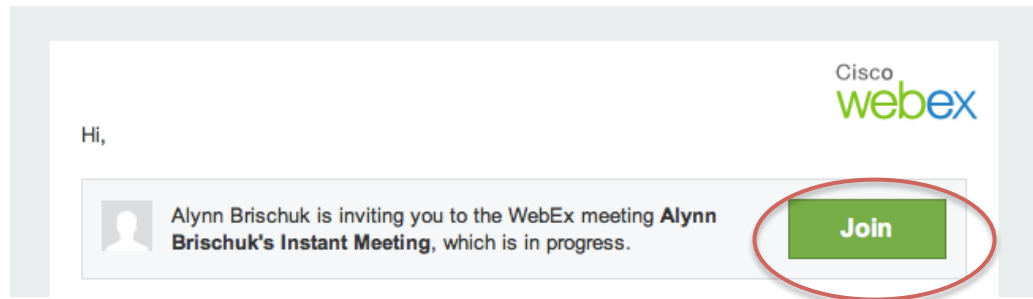
To complete Stage 1, the following steps will occur:

A) Meeting Setup:

- Please schedule a one hour time with Alynn to watch the video and complete the first rating by May 15th, 2013:
 - ❖ email alynn.brischuk@usask.ca, or call 1-306-370-6555
 - ❖ times are flexible and can be morning, afternoon or evening depending on the day
- You will receive an email confirmation of the meeting time, as well as a copy of the assessment form edited for easy ratings. The document will be called *STAGE 1 Video Rating Form.docm*. Please note that Microsoft Word is required as it is a Macros Enabled File. If this is a problem, please inform Alynn in your meeting setup.
- Review the form so you are familiar with it and save it to your computer hard drive. You will open this form at the time of the video observation.
- You will receive a  meeting invitation, a minimum of 24 hours prior to the meeting. If you do not see this meeting request, check your mailbox for junk mail just in case your mail program files this request inappropriately. If you still do not see a request, contact Alynn by email or phone.
- **You are not required to have a Webex account to join a Webex meeting.** No additional software downloads are needed, joining the meeting only requires you to Join from an email invitation. The meeting window will require your computer to be “Java” and “cookies” enabled to open. If you are interested in learning more about Webex meetings: <https://www.webex.com/login/join-meeting-tips>, and the [system requirements](#).
- A test meeting can be scheduled if you are unsure of your system requirements.

B) Starting the Meetings

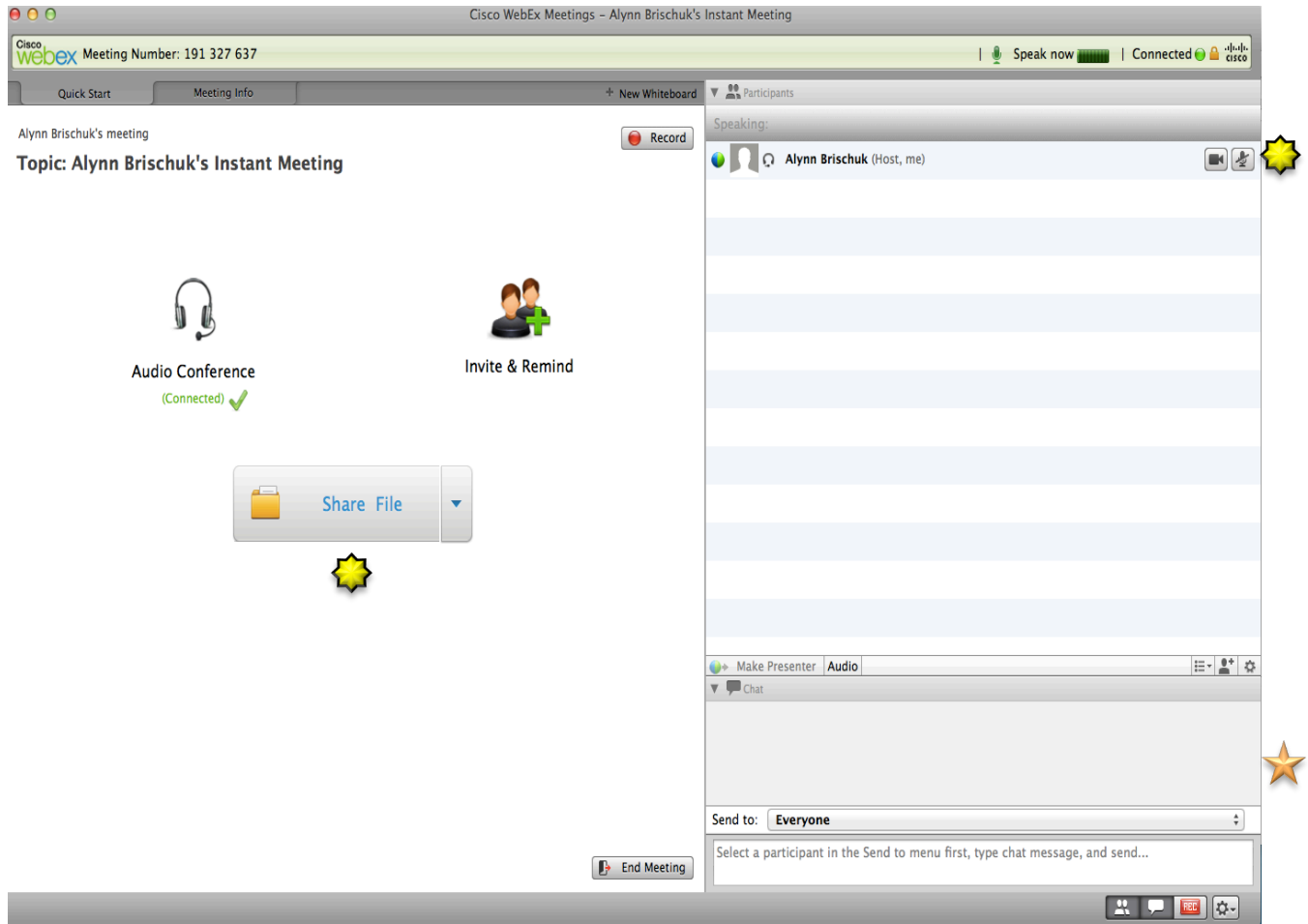
- At the planned meeting time, open the email invitation and **click “Join”** to enter the meeting.



- Try to ensure that no other windows or web browsing pages are open at the same time, this will allow for a better quality of the meeting and the video.
- A separate browser window will open the meeting space; it may take a few seconds to connect.
- You will be asked what audio input you will use for the meeting, options include to use a built in microphone in the computer, or to attach a headset:



- In the meeting, you may chat with Alynn via **webcam**★, or without video but with audio.
- There is also a **text only chat feature**★ if you do not have audio or video capability.



- If you have a webcam, you can control the display by clicking on your video camera icon.



- A webcam will be enabled when the icon looks like this:



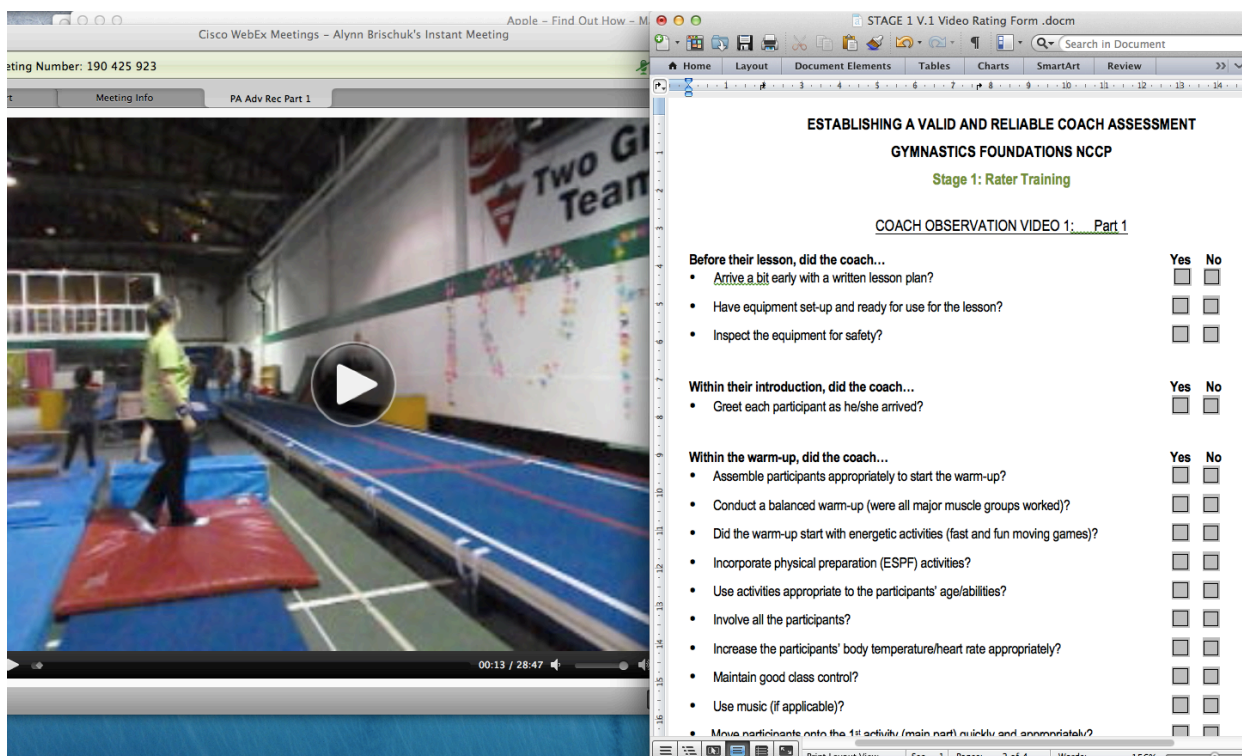
- Audio is muted when the icon looks like this:



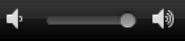

C) Video Observation and Coding

- Open your document, *STAGE 1 Video Rating Form.docm*. You can choose to setup the rating form and the video observation (internet) window to be visible at the same time, this option allows you to read and rate as the video is shown (*sample shown below).
- The video will be stopped in between sections so if you prefer, you may view the video on a full screen and at the completion of the section open the rating form.
- A third option is to print the document and follow on paper during the process and then transfer your ratings onto an electronic version afterwards if you prefer.
- Review the items on the form by reading them section by section. Identify any items you have questions about or are uncertain about.

**Sample view of two windows side by side to view the video and rate the criteria simultaneously:*



- The video clips will be played from the meeting host's computer (Alynn's). Only the meeting host can stop and start the video. Please communicate by the webcam or the chat features if you need a section to be re-played or stopped.

- The video will be stopped after each segment to allow rating, and will continue to the next segment when the rater indicates they are ready.
- On the video screen you have control over:
 - the volume 
 - the size of the window 
- A timer shows the point of the video and the total length.



Stage 1 Final Steps:

- Remember to **save** the rating form with your assessment to your hard drive, it is a good idea to save it after each video clip to not lose any of the ratings
- Email the assessment form once it is completed back to alynn.brischuk@usask.ca

If you have any questions please do not hesitate to contact us.

We sincerely thank you for your time and contribution to this valuable research,

Alynn and Keith

Alynn Brischuk
 MSc Candidate, College of Kinesiology
 University of Saskatchewan
 PAC, 87 Campus Drive
 Saskatoon SK S7N 5B2
alynn.brischuk@usask.ca
 P: 306-966-1123
 F: 306-966-6464

Dr. Keith Russell
 College of Kinesiology
 University of Saskatchewan
 PAC, 87 Campus Drive
 Saskatoon SK S7N 5B2
keith.russell@usask.ca
 P: 306-966-1067
 F: 306-966-6464

ESTABLISHING A VALID AND RELIABLE COACH ASSESSMENT

GYMNASTICS FOUNDATIONS NCCP

Stage 1: Rater Training

Who: 5 Experts randomly chosen

Instructions:

Please use the following sections from the Gymnastics Foundations Evaluation (2011, v.7) to complete your observation of the video clip shown.

If you see the coach has performed the behaviour, select the box for “Yes”. If you do not see the behaviour or believe the coach has adequately met the criteria as stated, select “No”.

Indicate your rating by clicking on the box under the rating you give, it is formatted to make an ☒ in the box for you.

Before their lesson, did the coach...

- Arrive a bit early with a written lesson plan?
- Have equipment set-up and ready for use for the lesson?
- Inspect the equipment for safety?

Yes No

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

Within their introduction, did the coach...

- Greet each participant as he/she arrived?

Yes No

<input type="checkbox"/>	<input type="checkbox"/>
--------------------------	--------------------------

Within the warm-up, did the coach...

- Assemble participants appropriately to start the warm-up?
- Conduct a balanced warm-up (were all major muscle groups worked)?
- Did the warm-up start with energetic activities (fast and fun moving games)?
- Incorporate physical preparation (ESPF) activities?
- Use activities appropriate to the participants' age/abilities?
- Involve all the participants?
- Increase the participants' body temperature/heart rate appropriately?
- Maintain good class control?

Yes No

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

- Use music (if applicable)? ☐ ☐
- Move participants onto the 1st activity (main part) quickly and appropriately? ☐ ☐

Within the main part of the lesson, did the coach...

- | | Yes | No |
|---|--------------------------|--------------------------|
| • Incorporate physical preparation (ESPF) and motor preparation activities? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Break skill learning into sequential steps? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Use drills / progressions / activities appropriate to the participants' age/abilities? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Use different teaching methods for participants with different learning styles? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Use a variety of drills / progressions / activities to ensure lots of little successes? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Recognize and adapt for an individual's needs? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Stress 1 to 3 key coaching points and safety? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Provide clear and concise instructions? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Correct errors as they occur by providing constructive & specific feedback? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Provide feedback to reinforce 'effort, trying, etc.'? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Have everyone involved through continuous activity? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Supervise all participants at all times? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Use supplementary equipment to assist with teaching / learning of skills? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Use equipment (large and small) effectively? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Do activities that work towards / achieve the stated objectives of the lesson? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Maintain class control? | <input type="checkbox"/> | <input type="checkbox"/> |

Within the cool-down, did the coach...

- | | Yes | No |
|---|--------------------------|--------------------------|
| • Assemble participants appropriately to start the cool-down? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Involve all participants? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Maintain good class control? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Do slower paced and/or stretching activities? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Finish off with a fun activity? | <input type="checkbox"/> | <input type="checkbox"/> |

Within the conclusion, did the coach...

- | | Yes | No |
|---|--------------------------|--------------------------|
| • Organize the participants appropriately to leave the gym? | <input type="checkbox"/> | <input type="checkbox"/> |

Overall communication: did the coach...

- | | Yes | No |
|--|--------------------------|--------------------------|
| • Catch the participants' attention quickly? | <input type="checkbox"/> | <input type="checkbox"/> |

- Provide clear and concise feedback and instructions? ☐ ☐
- Check the participants' understanding with regards to instruction and safety? ☐ ☐
- Maintain good, clear voice control? ☐ ☐
- Use non-verbal communication with the participants? ☐ ☐
- Address participants respectfully? ☐ ☐
- Have fun with the participants? ☐ ☐
- Handle disciplinary situations appropriately? ☐ ☐
- Handle any injury / emergency situation that occurred appropriately? ☐ ☐

General points - Personal: did the coach...

- | | Yes | No |
|---|--------------------------|--------------------------|
| • Dress appropriately for the lesson? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Act in a professional manner at all times? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Show enthusiasm? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Display a positive attitude towards their coaching? | <input type="checkbox"/> | <input type="checkbox"/> |

General points - Presentation: did the coach...

- | | Yes | No |
|---|--------------------------|--------------------------|
| • Present the right amount of material? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Present material in a logical way? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Present material at the right level for the participants? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Demonstrate a sound knowledge of the material being presented? | <input type="checkbox"/> | <input type="checkbox"/> |
| • Cover all key points / information relevant to the desired learning experience? | <input type="checkbox"/> | <input type="checkbox"/> |

APPENDIX E

Dear _____,

Thank you for your participation in Stage 1 of our investigation into *Establishing a Valid and Reliable Competency Assessment for Gymnastics Coach Certification*.

We originally anticipated doing a repeat testing in the training phase but we feel that the goals of the training are complete:

- a. orientation to the observation method and software
- b. orientation to coaching behaviours (evidences) listed in GCG 'evaluation form
- c. practice using the form

Stage 2: Content Validity Testing

Content Validity Testing:

Stage 2 is designed to establish "Content Validity". Each expert will provide input on the **content validity** of the behaviours in the draft evaluation form. Content Validity is: ***the extent to which the instrument measures what it is intended to measure*** (van der Mars, 1989).

To achieve content validity, your expert opinion is needed to assess the behaviours in the draft instrument for certification in ***NCCP Community Sport (Gymnastics Foundations), Artistic Gymnastics***.

Gymnastics Foundations Certification requires training in:

- Gymnastics Foundations Introduction – 2 day course
- Gymnastics Foundations Theory – 1 day course (Plan a Lesson and Making Ethical Decisions)
- Gymnastics Foundations Technical – 1 day course in a gymnastics discipline (Artistic)

The goal of the Gymnastics Foundations Evaluation is to verify the coach can "develop, implement and evaluate ***a safe, age and level-appropriate gymnastics lesson***" (Gymnastics Canada Gymnastique, 2010). Evaluation of coaching competency will include a portfolio completed by the coach and includes a coach profile, self-assessment, session plan, lesson plans, an emergency action plan, and a formal observation currently suggested to be done on video/DVD.

Evaluation for NCCP Certification (NCCP Evaluation Toolkit v.2.1, 2011):

- Is evidence-based, meaning that relevant information is collected to make a decision
- Is preceded by NCCP training or relevant experience
- Consists of observable coaching behaviour that is used to determine whether a coach meets a given criteria
- Recognizes and respects individual coaching styles
- Has requirements, procedures and methods that are: administratively feasible, professionally acceptable, publicly credible, legally defensible, economically affordable, and reasonably accessible

Formal Observation

- Coaching Association of Canada requires formal observation in the certification process
- Observation helps determine the coaches ability to demonstrate specific coaching outcomes
- The coach's behaviours/actions provide observable evidence related to the outcomes
- Evidence is gained at the time of coaching
- Observation requires the coach to be working in the appropriate sport context
- Observation allows for feedback, safety issues, and action planning to be discussed between the evaluator and the coach

To complete stage 2, the following steps will occur:

- You will receive an e-mail with a **Content Validity Index** (CVI) survey. The document is titled: *Stage 2 Content Validity Index Final.docm*. It is a Microsoft Word Macro enabled document.
 - The CVI is a survey on a 4 point scale ranging from “1 – the behaviour is not essential” to “4- the behaviour is very essential”
- On your own time, please review all behaviours in the CVI and think thoroughly about its importance to Gymnastics Foundations (Artistic) coaching.
- Complete the survey by rating each behaviour on the index scale of how essential you feel it is.
- Room is provided for comments on behaviours that you rate 1 “not essential” or 2 “somewhat essential”. Any feedback on why you feel it is not essential is valuable, along with what changes you would suggest for it to become essential (if anything).
- Two summary questions are open-ended to find out if you feel the instrument is missing any essential elements for coaching in this context. A second question seeks your input on the scoring system.
- You can complete this on your own time. It should take no more than one hour. There is room for comments on changes to the wording of the behaviour criteria.
- Please try to complete the form within one week of receiving it. If it is not returned within one week, a reminder e-mail will be sent from the research team to inquire if you are able to complete it.

Stage 2 Final Steps:

- Remember to **save** the rating form with your assessment to your hard drive, it is a good idea to save it after each section
- Email the completed survey to alynn.brischuk@usask.ca

Expert opinion is the primary form of content validity testing and therefore your input is considered valuable in understanding potential changes or accuracy of the behaviours sought in the coaching observation/assessment.

We thank you for your time and expertise in this stage.

STAGE 2. CONTENT VALIDITY INDEX
GYMNASTICS FOUNDATIONS NCCP EVALUATION

Part A Instructions:

As an expert in the Gymnastics Foundations NCCP, you will determine the validity of the behaviours in the Gymnastics Foundations Evaluation.

Please consider the **context-specific environment** that Gymnastics Foundations Coaches work in, and carefully consider how essential each of the behaviours are to being a competent coach.

The **Content Validity Index** is a four-point “Likert” Scale:

- 1 = the behaviour is not essential
- 2 = the behaviour is somewhat essential, it would need revisions to become quite or very essential
- 3 = the behaviour is quite essential or relevant. It may need minor revision.
- 4 = the behaviour is very essential

Indicate your opinion by marking an ☒ in the box under the rating you feel strongest about.

If you provide a rating of 1 or 2, please provide comments or suggestions as to why, or what revisions you would consider appropriate for the behaviour to become a 3 “quite essential” or 4 “very essential”.

1 = not essential 2= somewhat essential 3 = quite essential 4 = very essential

Before their lesson, did the coach...

- | | 1 | 2 | 3 | 4 |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| • Arrive a bit early with a written lesson plan? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Have equipment set-up and ready for use for the lesson? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Inspect the equipment for safety? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Within their introduction, did the coach...

- | | 1 | 2 | 3 | 4 |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| • Greet each participant as he/she arrived? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Comments for before lesson and INTRODUCTION behaviours (IF Rated 1 or 2):

1 = not essential 2= somewhat essential 3 = quite essential 4 = very essential

Within the warm-up, did the coach...

	1	2	3	4
• Assemble participants appropriately to start the warm-up?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Conduct a balanced warm-up (were all major muscle groups worked)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Did the warm-up start with energetic activities (fast and fun moving games)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Incorporate physical preparation (ESPF) activities?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Use activities appropriate to the participants' age/abilities?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Involve all the participants?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Increase the participants' body temperature/heart rate appropriately?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Maintain good class control?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Use music (if applicable)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Move participants onto the 1 st activity (main part) quickly and appropriately?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comments for WARM-UP behaviours (IF Rated 1 or 2):

Within the main part of the lesson, did the coach...

	1	2	3	4
• Incorporate physical preparation (ESPF) and motor preparation activities?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Break skill learning into sequential steps?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Use drills / progressions / activities appropriate to the participants' age/abilities?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Use different teaching methods for participants with different learning styles?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Use a variety of drills / progressions / activities to ensure lots of little successes?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Recognize and adapt for an individual's needs?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Stress 1 to 3 key coaching points and safety?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Provide clear and concise instructions?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Correct errors as they occur by providing constructive & specific feedback?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Provide feedback to reinforce 'effort, trying, etc.'?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Have everyone involved through continuous activity?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Supervise all participants at all times?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

1 = not essential 2= somewhat essential 3 = quite essential 4 = very essential

- | | 1 | 2 | 3 | 4 |
|--|--------------------------|--------------------------|--------------------------|--------------------------|
| • Use supplementary equipment to assist with teaching / learning of skills? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Use equipment (large and small) effectively? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Do activities that work towards / achieve the stated objectives of the lesson? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Maintain class control? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Comments for MAIN PART behaviours (IF Rated 1 or 2):

Within the cool-down, did the coach...

- | | 1 | 2 | 3 | 4 |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| • Assemble participants appropriately to start the cool-down? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Involve all participants? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Maintain good class control? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Do slower paced and/or stretching activities? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Finish off with a fun activity? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Within the conclusion, did the coach...

- | | 1 | 2 | 3 | 4 |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| • Organize the participants appropriately to leave the gym? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Comments for COOL-DOWN AND CONCLUSION behaviours (IF Rated 1 or 2):

Overall communication: did the coach...

- | | 1 | 2 | 3 | 4 |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| • Catch the participants' attention quickly? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Provide clear and concise feedback and instructions? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Check the participants' understanding with regards to instruction and safety? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Maintain good, clear voice control? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Use non-verbal communication with the participants? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Address participants respectfully? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

1 = not essential 2= somewhat essential 3 = quite essential 4 = very essential

- | | 1 | 2 | 3 | 4 |
|--|--------------------------|--------------------------|--------------------------|--------------------------|
| • Have fun with the participants? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Handle disciplinary situations appropriately? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Handle any injury / emergency situation that occurred appropriately? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

General points - Personal: did the coach...

- | | 1 | 2 | 3 | 4 |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| • Dress appropriately for the lesson? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Act in a professional manner at all times? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Show enthusiasm? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Display a positive attitude towards their coaching? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

General points - Presentation: did the coach...

- | | 1 | 2 | 3 | 4 |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| • Present the right amount of material? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Present material in a logical way? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Present material at the right level for the participants? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Demonstrate a sound knowledge of the material being presented? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| • Cover all key points / information relevant to the desired learning experience? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Comments for OVERALL behaviours (IF Rated 1 or 2):

Part B - Please answer the questions in as much detail as you can:

a. Are any behaviours omitted that are essential to coaching Artistic Gymnastics in NCCP Community Sport?

b. Does the scoring rubric (YES/NO selection) appropriately measure the behaviours?
Why or why not?

A rubric is understood as an assessment tool that describes a level of performance on a task - in this case Competent / Not Competent. Keep in mind that a scoring rubric should describe general criteria that can be witnessed across individual performances but may not account for unique personal characteristics.

When you are satisfied with your survey answers, please return an electronic copy via email, or fax a paper copy to Alynn Brischuk, alynn.brischuk@usask.ca

APPENDIX F

ESTABLISHING A VALID AND RELIABLE COACH ASSESSMENT

GYMNASTICS FOUNDATIONS NCCP

Modified Instrument July 2013

Gymnastics Lesson Observation Form

Coach's name:				CC number:			
Address:							
	Street including apartment number	City	P/T	Postal Code			
Phone:				Email:			
Supervisor's name:							
Club:				Location:			
		City	P/T				
Phone:				Email:			

The following tasks are observed and categorized from *Insufficient* to *Excellent*. You are encouraged to use the ratings and the comments as feedback on your strengths and weaknesses of your coaching.

Excellent - The coach performs the task in an outstanding way; exceptional coaching that reflects mastery of the task.

Good - The coach performs the task successfully. The actions taken are appropriate, though a few opportunities were missed that can be improved upon to reach mastery.

Fair - The coach attempts the task but is missing some elements of a successful performance.

Insufficient - The coach attempts the task but requires significant improvement to be successful. Or the coach did not provide evidence of the task (not shown).

The results of the ratings will help the assessor make an overall competency assessment, page 5.

GYMNASTICS FOUNDATIONS OBSERVATION

Before their lesson, did the coach...	<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
Arrive early?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dress appropriately for the lesson?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Setup equipment to be available, ready and safe for use in the lesson?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Greet each participant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comments: _____

Within the Warm-Up, did the coach...	<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
Assemble participants appropriately to start the warm-up?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Conduct a balanced warm-up (were all major muscle groups worked)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Begin the lesson with enjoyable and energetic activities to increase participant body temperature?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Incorporate activities to develop physical components (such as Endurance, Strength, Power or Flexibility)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use activities appropriate to the participants' age/abilities?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Involve all the participants?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Maintain good class control?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Move participants onto the 1 st activity (main part) quickly and appropriately?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comments: _____

Within the Main Part of the lesson, did the coach...	<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
Provide clear and concise instructions?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Teach skills by sequential steps (drills/progressions)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Teach to the participants' age and abilities?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Teach to ensure many little successes?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Recognize and adapt for an individual's needs?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Correct errors as they occur by providing constructive & specific feedback?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Provide feedback to reinforce effort, trying, etc.?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Keep participants moving through continuous activity (such as stations/circuits)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Supervise the participants within their view at all times?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Use supplementary equipment to assist with teaching / learning of skills?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use equipment (large and small) effectively?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Incorporate activities to develop both physical (ESPF) and motor components (ABCS)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Follow their plan and use activities that match their objectives of the lesson?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Maintain good class control?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comments: _____

Within the Cool-Down, did the coach...	<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
Assemble participants appropriately to start the cool-down?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Involve all participants?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Maintain good class control?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
End the class with an enjoyable activity?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Within the Conclusion, did the coach...	<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
End the class on a positive note?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comments: _____

Overall communication: did the coach...	<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
Catch the participants' attention quickly?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Provide clear and concise feedback and instructions?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ask if the participants understand or if they have questions?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Maintain good, clear voice control?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use non-verbal communication with the participants?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Address participants with respect and kindness?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Have fun with the participants (ex.use humour)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Handle any unplanned situations appropriately such as equipment/safety/disciplinary/emergency)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comments: _____

General points - Personal: did the coach...	<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
Act in a professional manner at all times?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Show enthusiasm?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Display a positive attitude towards their coaching?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

General points - Presentation: did the coach...	<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
Present the right amount of material?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Present material in a logical way?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Present material at the right level for the participants?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Demonstrate a sound knowledge of the material being presented?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cover all key points / information relevant to the desired learning experience?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comments: _____

COACHING COMPETENCY ASSESSMENT RESULTS

For Certification of Coaching Gymnastics Foundations, the coach is evaluated on their ability to implement a safe, age and level-appropriate gymnastics lesson.

Expert	Coach demonstrates an extensive base of knowledge and expertise through perfectly meaningful activities and a naturally superior performance for the age and level of the athletes.	<input type="checkbox"/>	<p style="text-align: center;">COMPLETE</p> <p style="text-align: center;"><i>The observation portion of the evaluation is successfully completed.</i></p>	COMPLETE
Proficient	Coach displays intuitive control of a safe, age and level appropriate lesson and uses sophisticated instructional techniques that focus on the most critical components for individual athlete success.	<input type="checkbox"/>		
Competent	Coach implements the expected procedures of a safe, age and level appropriate lesson, while displaying a desire to see athletes learn, develop and grow.	<input type="checkbox"/>		
Beginner	Coach requires more knowledge and experience to successfully implement a safe, age and level appropriate lesson.	<p style="text-align: center;">INCOMPLETE</p> <ul style="list-style-type: none"> <input type="checkbox"/> The lesson shows that the coach is not ready to be certified. <input type="checkbox"/> The lesson has issues concerning safety. <input type="checkbox"/> The lesson shows harmful or unethical coaching behaviours. <p style="text-align: center;"><i>The evaluator has provided feedback and will discuss action plan items to lead to your future success.</i></p>		INCOMPLETE

Comments: _____

Action Items: _____

Evaluator's signature: _____ **Date:** _____

Coach signature: _____ **Date:** _____

APPENDIX G

July 2013,



Dear _____,

Thank you for your commitment to participate in our study "*Establishing a Valid and Reliable Competency Assessment for Gymnastics Coach Certification*". As part of your consent to participate in our study, we informed you of the opportunity to provide your input on the validity of the evaluation criteria for Gymnastics Foundations coaches.

Your input is highly valued to understanding what behaviours are accurate, appropriate, and best describe coaching competency in the NCCP for Artistic Gymnastics Foundations. We are interested in determining if the evaluation criteria for coaches in this context has **Face Validity**, which is a measure of how relevant an instrument is at assessing job tasks from the view of people who are experienced in the job (Furr & Bacharach, 2008).

The research survey that we would like you to complete is directly corresponding to the modified Gymnastics Foundations Evaluation, originally drafted by Gymnastics Canada and since reviewed and modified by an expert group through our study.

As per the consent to participate, your input will remain anonymous.

About Competency:

- Competency refers to a person's ability to perform a certain role or a set of desired behaviours (such as in a job)
- Behaviours are observable actions based on a person's knowledge, skills, personal characteristics, and aptitude (Lucia & Lepsinger, 1999)
- A competent performer is beyond a beginner stage, but is not yet known to be proficient or expert in their field
- Competency assessment must be reflective of the work required in the context of the person's work

About NCCP for Artistic Gymnastics Foundations:

- Gymnastics Foundations is the NCCP training and certification for Community Sport Coaches offered by Gymnastics Canada
- The goal of the Gymnastics Foundations Evaluation is to verify the coach can "develop, implement and evaluate a safe, age and level-appropriate gymnastics lesson" (Gymnastics Canada Gymnastique, 2010).
- Evaluation of coaching competency will include a portfolio completed by the coach and includes a coach profile, self-assessment, session plan, lesson plans, an emergency action plan, and a formal observation currently suggested to be done on video/DVD.

Gymnastics Foundations Certification requires training in:

- Gymnastics Foundations Introduction – 2 day course
- Gymnastics Foundations Theory – 1 day course (Plan a Lesson and Making Ethical Decisions)
- Gymnastics Foundations Technical – 1 day course in a gymnastics discipline (Artistic)

To provide your input, please complete the following steps. The survey should take no more than 1 hour of your time.

- A) Open the attached survey
 - the document is titled *Stage 3 Face Validity Survey.doc*
- B) Open the reference document
 - the survey questions are related to *Modified Assessment July 2013.pdf*
- C) Read the instructions and review the assessment tasks from the perspective of a coach who would receive an evaluation
- D) Complete the survey questions to the degree you feel comfortable
- E) Final Steps:
 - Remember to **save** the survey with your comments, it is a good idea to save it after each section.
 - Email the survey when you are happy with your responses to alynn.brischuk@usask.ca

If you have any questions please do not hesitate to contact us.

We sincerely thank you for your time and contribution to this valuable research,

Alynn and Keith

Alynn Brischuk
MSc Candidate, College of Kinesiology
University of Saskatchewan
PAC, 87 Campus Drive
Saskatoon SK S7N 5B2
alynn.brischuk@usask.ca
P: 306-966-1123
F: 306-966-6464

Dr. Keith Russell
College of Kinesiology
University of Saskatchewan
PAC, 87 Campus Drive
Saskatoon SK S7N 5B2
keith.russell@usask.ca
P: 306-966-1067
F: 306-966-6464

ESTABLISHING A VALID AND RELIABLE COACH ASSESSMENT

GYMNASTICS FOUNDATIONS NCCP

Stage 3: Face Validity

Who: 4 Active Certified Level 1 Coaches, minimum 5 years of experience

What: Face Validity is the level of relevance that the tasks in the assessment reflect the actual duties and job tasks from the viewpoint of individuals who are experienced with performing in the coaching context.

Stage Goal: Establish Face Validity of the NCCP Artistic Gymnastics Foundations Coaching Assessment

Stage 3 Instructions:

Please use the document Gymnastics Foundations Assessment Form modified July 2013 (PDF format) to complete the questions below.

Consider the relevance of the tasks to Gymnastics Foundations (Community Sport) Artistic programming, specifically targeting gymnastics participants 6 years and older in recreational classes. Coach competency for this context of coaching is determined by the coach's ability **to implement a safe, age and level-appropriate gymnastics lesson (Gymnastics Canada, 2011).**

Use the grey highlighted text boxes to answer the survey questions. Please note that your answers will be coded to keep your anonymity and confidentiality of your participation in this project.

The questions are:

Are all elements of the instrument reflective of coaching in the appropriate context (artistic gymnastics/community sport)?

Are important elements omitted? If so, which ones?

Are unimportant elements incorrectly included? If so, which ones?

When you have completed the survey questions, please follow the Survey Completion Actions.

ESTABLISHING A VALID AND RELIABLE COACH ASSESSMENT

GYMNASTICS FOUNDATIONS NCCP

Stage 3: Face Validity

SURVEY QUESTIONS

1. Are all elements of the instrument reflective of coaching in the appropriate context (artistic gymnastics/community sport)?

Type your answer in shaded box:

2. Are important elements omitted? If so, which ones?

3. Are unimportant elements incorrectly included? If so, which ones?

STAGE 3 COMPLETION ACTIONS

1. Please SAVE this document with your comments to your hard drive.
2. Complete the honoraria requisition form with your address so we can mail you a thank-you
3. Email both documents, *STAGE 3 Face Validity Survey* and *Honoraria Requisition* to alynn.brischuk@usask.ca

Thank you for your time.

APPENDIX H

Dear

First we would like to apologize for the lengthy delay in contacting you after your expressed interest to participate in our study, *“Establishing a Valid and Reliable Competency Assessment for Gymnastics Coach Certification”*.

We received a good response from the Learning Facilitator group that we contacted, and as a result we randomized the group into two participant groups – one that would complete Stages 1 and 2, and one that would called upon for Stages 4 and 5.

We are now complete the first 3 stages of the 5 stage process to establish validity and reliability. Each stage has lead to interesting findings (and unqiue challenges) and we are now prepared to run the final two stages with your expert assistance.

Stage 4 is a crucial stage, where we will test the inter-rater reliability of the coaching assessment instrument.

If you are still interested in participating, we are glad to provide you with a small compensation for your time as a funded project through the Coaching Association of Canada's *Coach Education Research Grant*.

I have attached a detailed letter with the information you need to participate. The reliability test will require a computer and an internet connection.

If you have any questions following the details in this document on how to proceed, please do not hesitate to contact me.

I'm looking forward to working with you in establishing the reliability of such an integral piece of our NCCP program!

Sincerely,
Alynn

Alynn Brischuk
MSc Candidate, College of Kinesiology
PAC, 87 Campus Drive
Saskatoon SK S7N 5B2
alynn.brischuk@usask.ca
P: 306-966-1123
C: 306-370-6555

This research is done with support from:



Dear _____,

Thank you for your involvement in our study "*Establishing a Valid and Reliable Competency Assessment for Gymnastics Coach Certification*". We are now complete the first three stages of the five stage research process. Each stage has lead to interesting findings (and unique challenges) and we are now prepared to run the final two stages.

Your expertise is once again called upon for Stage 4, and potentially Stage 5 if you are randomly selected. Stage 4 requires 10 expert participants, while Stage 5 will be a selection of half of this group to run a repeat test.

In Stage 1, through video observation and coding, we trained a few experts on our process and determined the initial reliability levels of a competency assessment instrument. We had hoped to have all experts above 0.80 Percent Agreement and 0.61 Agreement Coefficient, but what we found was a very large range of agreement.

The results of Stage 1 were:

Between Each Expert and a Master Coder:

Percent Agreement Range = 0.54 to 0.95 (54% to 95%)

Agreement Coefficient (Accounts for chance) Range = 0.24 to 0.96 (Fair to Almost Perfect)

Results Comparing 5 Expert Raters:

Percent Agreement = 0.70 (70%)

Agreement Coefficient = 0.56 (Moderate)

Inter-Rater Paired Comparisons:

Highest Percent Agreement = 0.96; Higher AC = 0.9

Lowest Percent Agreement = 0.5; Lowest AC = 0.15

We also discovered that having more items to code (breaking out the main part into 3 apparatus) created significantly less reliability:

Percent Agreement = 0.51 to 0.85 (51% to 85%)

Agreement Coefficient (Accounts for chance) = 0.06 to 0.8 (Slight to Substantial)

To Participate in Stages 4 and 5:

I have attached a detailed letter with the information you need to participate. Stage 4 will be a repeat of the process we ran for Stage 1, as a video observation online using Webex conferencing tools.

If you are still interested in participating, we are glad to provide you with another small compensation for your time in each stage through the funding from Coaching Association of Canada's *Coach Education Research Grant*.

If you have any questions following the details in this document on how to proceed, please do not hesitate to contact me.
I'm looking forward to working with you in establishing the reliability of such an integral piece of our NCCP program!

Sincerely,
Alynn

Alynn Brischuk
MSc Candidate, College of Kinesiology
PAC, 87 Campus Drive
Saskatoon SK S7N 5B2
alynn.brischuk@usask.ca
P: 306-966-1123
C: 306-370-6555



This research is done with support from:



September 2013,

Dear Research Participant,


Thank you for your commitment to participate in our study *“Establishing a Valid and Reliable Competency Assessment for Gymnastics Coach Certification”*.

If you have any questions following the details in this document on how to proceed with the research stage, please do not hesitate to contact us.

Stage 4 is designed to test Inter-Rater Reliability. Each rater will independently view and rate the behaviours they see in the video to the coach assessment form. The assessment form to be used in training is a modified version of the Gymnastics Foundations Evaluation.

To complete Stage 4, the following steps will occur:

A) Meeting Setup:

- Please schedule a one hour time with Alynn to watch the video and complete the first rating by October 15 2013:
 - ❖ email alynn.brischuk@usask.ca, or call 1-306-370-6555
 - ❖ times are flexible and can be morning, afternoon or evening depending on the day
- You will receive an email confirmation of the meeting time, as well as a copy of the assessment form edited for easy ratings. The document will be called *Video Rating Form Stage 4.docm*. Please note that Microsoft Word is required as it is a Macros Enabled File. If this is a problem, please inform Alynn in your meeting setup.
- Review the form so you are familiar with it and save it to your computer hard drive. You will open this form at the time of the video observation.
- You will receive a  meeting invitation, a minimum of 24 hours prior to the meeting. If you do not see this meeting request, check your mailbox for junk mail just in case your mail program files this request inappropriately. If you still do not see a request, contact Alynn by email or phone.
- **You are not required to have a Webex account to join a Webex meeting.** No additional software downloads are needed, joining the meeting only requires you to Join from an email invitation. The meeting window will require your computer to be “Java” and “cookies” enabled to open. If you are interested in learning more about Webex meetings: <https://www.webex.com/login/join-meeting-tips>, and the [system requirements](#).
- A test meeting can be scheduled if you are unsure of your system requirements.

B) Starting the Meetings

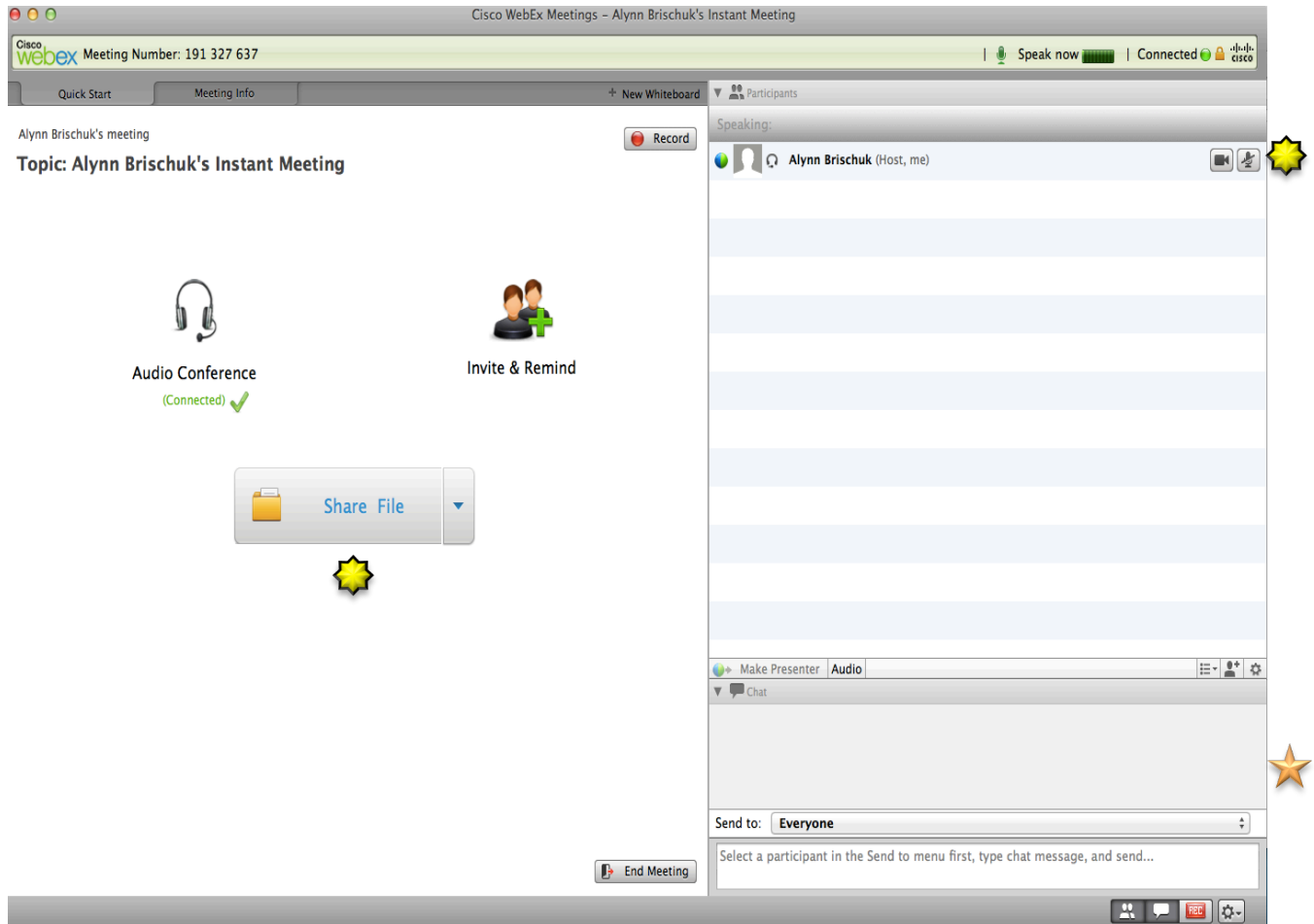
- At the planned meeting time, open the email invitation and **click “Join”** to enter the meeting.



- Try to ensure that no other windows or web browsing pages are open at the same time, this will allow for a better quality of the meeting and the video.
- A separate browser window will open the meeting space; it may take a few seconds to connect.
- You will be asked what audio input you will use for the meeting, options include to use a built in microphone in the computer, or to attach a headset:



- In the meeting, you may chat with Alynn via **webcam**★, or without video but with audio.
- There is also a **text only chat feature**★ if you do not have audio or video capability.



- If you have a webcam, you can control the display by clicking on your video camera icon.



- A webcam will be enabled when the icon looks like this:



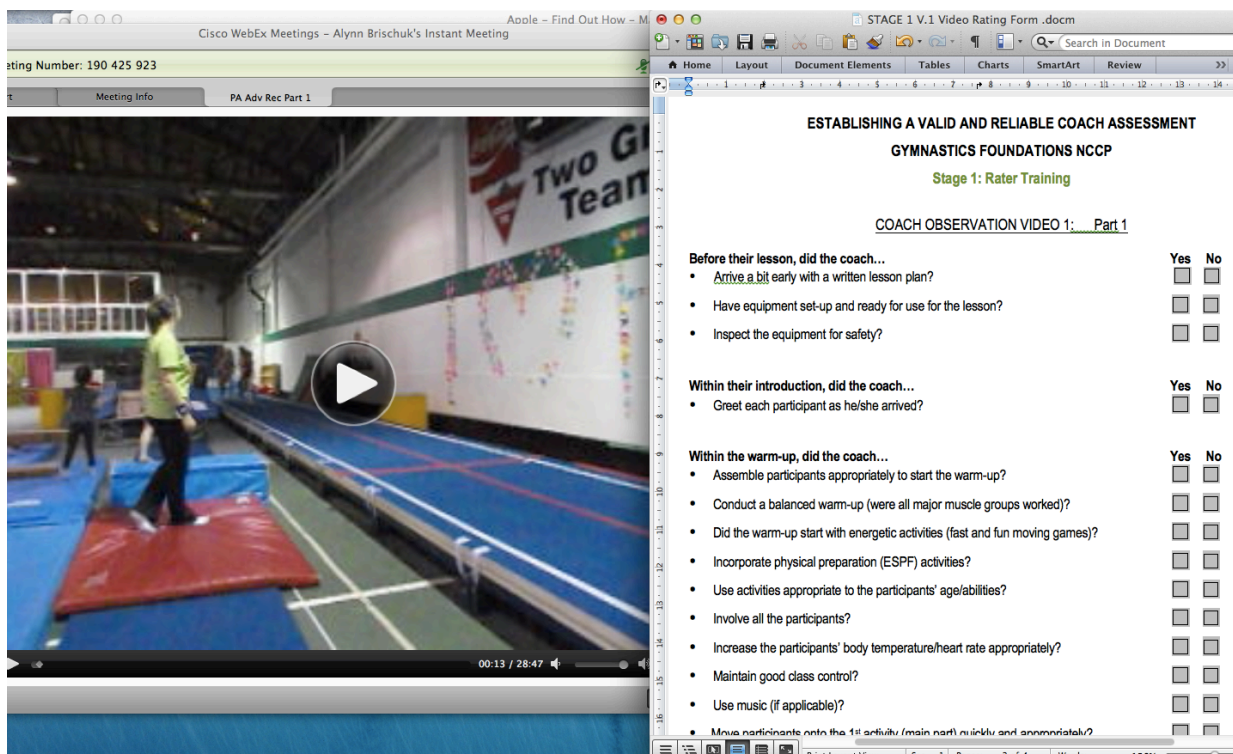
- Audio is muted when the icon looks like this:





C) Video Observation and Coding

- Open your document, *Video Rating Form Stage 4.docm*. You can choose to setup the rating form and the video observation (internet) window to be visible at the same time, this option allows you to read and rate as the video is shown (*sample shown below).
- The video will be stopped in between sections so if you prefer, you may view the video on a full screen and at the completion of the section open the rating form.
- A third option is to print the document and follow on paper during the process and then transfer your ratings onto an electronic version afterwards if you prefer.
- Review the items on the form by reading them section by section. Identify any items you have questions about or are uncertain about.

**Sample view of two windows side by side to view the video and rate the criteria simultaneously:*



- The video clips will be played from the meeting host's computer (Alynn's). Only the meeting host can stop and start the video. Please communicate by the webcam or the chat features if you need a section to be re-played or stopped.
- The video will be stopped after each segment to allow rating, and will continue to the next segment when the rater indicates they are ready

- On the video screen you have control over:
 - the volume 
 - the size of the window 
- A timer shows the point of the video and the total length.



Stage 4 Final Steps:

- Remember to **save** the rating form with your assessment to your hard drive, it is a good idea to save it after each video clip to not lose any of the ratings
- Email the assessment form once it is completed back to alynn.brischuk@usask.ca

If you have any questions please do not hesitate to contact us.

We sincerely thank you for your time and contribution to this valuable research,

Alynn and Keith

Alynn Brischuk
MSc Candidate, College of Kinesiology
University of Saskatchewan
PAC, 87 Campus Drive
Saskatoon SK S7N 5B2
alynn.brischuk@usask.ca
P: 306-966-1123
F: 306-966-6464

Dr. Keith Russell
College of Kinesiology
University of Saskatchewan
PAC, 87 Campus Drive
Saskatoon SK S7N 5B2
keith.russell@usask.ca
P: 306-966-1067
F: 306-966-6464

APPENDIX I

ESTABLISHING A VALID AND RELIABLE COACH ASSESSMENT

GYMNASTICS FOUNDATIONS NCCP

Modified Instrument – Stage 4 & 5

Gymnastics Lesson Observation Form

Coach's name: _____ CC number: _____

Address: _____
Street including apartment number City P/T Postal Code

Phone: _____ Email: _____

Supervisor's name: _____

Club: _____ Location: _____
City P/T

Phone: _____ Email: _____

The following tasks are observed and categorized from *Insufficient* to *Excellent*. You are encouraged to use the ratings and the comments as feedback on your strengths and weaknesses of your coaching.

Excellent - The coach performs the task in an outstanding way; exceptional coaching that reflects mastery of the task.

Good - The coach performs the task successfully. The actions taken are appropriate, though a few opportunities were missed that can be improved upon to reach mastery.

Fair - The coach attempts the task but is missing some elements of a successful performance.

Insufficient - The coach attempts the task but requires significant improvement to be successful. Or the coach did not provide evidence of the task (not shown).

The results of the ratings will help the assessor make an overall competency assessment, page 5.

COACH OBSERVATION ASSESSMENT – TASKS & RATINGS

Before their lesson, did the coach...		<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
1	Arrive early?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	Dress appropriately for the lesson?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Setup equipment to be available, ready and safe for use in the lesson?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	Greet each participant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Task # Comment(s):

General Comment(s):

Within the Warm-Up, did the coach...		<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
5	Assemble participants appropriately to start the warm-up?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	Conduct a balanced warm-up (were all major muscle groups worked)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	Begin the lesson with enjoyable and energetic activities to increase participant body temperature?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	Incorporate activities to develop physical components (such as Endurance, Strength, Power or Flexibility)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	Use activities appropriate to the participants' age/abilities?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	Involve all the participants?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11	Maintain good class control?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12	Move participants onto the 1 st activity (main part) quickly and appropriately?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Task # Comment(s):

General Comment(s):

Within the Main Part of the lesson, did the coach...		<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
13	Provide clear and concise instructions?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14	Teach skills by sequential steps (drills/progressions)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15	Teach to the participants' age and abilities?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16	Teach to ensure many little successes?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17	Recognize and adapt for an individual's needs?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18	Correct errors as they occur by providing constructive & specific feedback?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19	Provide feedback to reinforce effort, trying, etc.?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20	Keep participants moving through continuous activity (such as stations/circuits)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21	Supervise the participants within their view at all times?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22	Use supplementary equipment to assist with teaching / learning of skills?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23	Use equipment (large and small) effectively?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24	Incorporate activities to develop both physical (ESPF) and motor components (ABCS)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25	Follow their plan and use activities that match their objectives of the lesson?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26	Maintain good class control?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Task # Comment(s):

General Comment(s):

Within the Cool-Down, did the coach...		<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
27	Assemble participants appropriately to start the cool-down?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28	Involve all participants?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29	Maintain good class control?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30	End the class with an enjoyable activity?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Within the Conclusion, did the coach...		<i>Excellent</i>	<i>Good</i>	<i>Fair</i>	<i>Insufficient</i>
31	End the class on a positive note?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Task # Comment(s):

General Comment(s):

Overall communication: did the coach...		Excellent	Good	Fair	Insufficient
32	Catch the participants' attention quickly?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33	Provide clear and concise feedback and instructions?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34	Ask if the participants understand or if they have questions?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
35	Maintain good, clear voice control?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
36	Use non-verbal communication with the participants?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
37	Address participants with respect and kindness?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
38	Have fun with the participants (ex.use humour)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
39	Handle any unplanned situations appropriately? (could be any of equipment/ lack of athlete interest / safety / disciplinary / emergency)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Task # Comment(s):

General Comment(s):

General points - Personal: did the coach...		Excellent	Good	Fair	Insufficient
40	Act in a professional manner at all times?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
41	Show enthusiasm?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
42	Display a positive attitude towards their coaching?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

General points - Presentation: did the coach...		Excellent	Good	Fair	Insufficient
43	Present the right amount of material?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
44	Present material in a logical way?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
45	Present material at the right level for the participants?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
46	Demonstrate a sound knowledge of the material being presented?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
47	Cover all key points / information relevant to the desired learning experience?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Task # Comment(s):

General Comment(s):

COACHING COMPETENCY ASSESSMENT RESULTS

For Certification of Coaching Gymnastics Foundations, the coach is evaluated on their ability to implement a safe, age and level-appropriate gymnastics lesson.

Expert	Coach demonstrates an extensive base of knowledge and expertise through perfectly meaningful activities and a naturally superior performance for the age and level of the athletes.	<input type="checkbox"/>	<p style="text-align: center;">COMPLETE</p> <p style="text-align: center;"><i>The observation portion of the evaluation is successfully completed.</i></p>	COMPLETE
Proficient	Coach displays intuitive control of a safe, age and level appropriate lesson and uses sophisticated instructional techniques that focus on the most critical components for individual athlete success.	<input type="checkbox"/>		
Competent	Coach implements the expected procedures of a safe, age and level appropriate lesson, while displaying a desire to see athletes learn, develop and grow.	<input type="checkbox"/>		
Beginner	Coach requires more knowledge and experience to successfully implement a safe, age and level appropriate lesson.	<p style="text-align: center;">INCOMPLETE</p> <p><input type="checkbox"/> The lesson shows that the coach is not ready to be certified.</p> <p><input type="checkbox"/> The lesson has issues concerning safety.</p> <p><input type="checkbox"/> The lesson shows harmful or unethical coaching behaviours.</p> <p><i>The evaluator has provided feedback and will discuss action plan items to lead to your future success.</i></p>		INCOMPLETE

Comments:

Action Items:

Evaluator's signature: _____ Date: _____

Coach signature: _____ Date: _____