University of Alberta

Development of a General Measure of Physical Self-Concept of Muscularity

by

Christina Candice Loitz

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

Doctor of Philosophy

Faculty of Physical Education and Recreation

Fall 2012
Edmonton, Alberta

Abstract

The purposes of this dissertation were to (a) develop an instrument that assess the physical self-concept of muscularity among young adults, (b) produce content validity, internal structural validity and external criterion-related validity evidence to support the new instrument, and (c) develop internal consistency and stability evidence supporting the new instrument. Physical self-concept and body image research generally focuses on body-fatness, but since the 1990s, perceptions of muscularity have been a growing concern for some men and women (McCreary & Sadava, 2001; McCreary & Sasse, 2000; Olivardia et al., 2004; Pope et al., 2000). Currently, self-concept of muscularity measures use silhouettes that show inconsistent validity and reliability (Cafri & Thompson, 2004). The two most popular physical self-concept instruments (Physical Self-Perception Profile, Fox & Corbin, 1989; and Physical Self-Description Questionnaire, Marsh, Richards, Johnson, Roche, & Tremayne,1994) do not include a measure of muscularity. The first study in this dissertation assessed the content relevance and representativeness of the new instrument based on ratings provided by an expert panel of academics. The second study assessed the content relevance and representativeness of the new instrument based on ratings provided by a panel of fitness experts. The third study assessed the internal structural validity of the instrument using exploratory factor analysis. The fourth study assessed the internal structural validity using exploratory factor analysis, criterion related validity (external variables included body mass index, physical activity level, exercise participation, exercise identity, drive for muscularity and drive for thinness) and internal consistency reliability and stability over a one-week time

period. The validity evidence presented in this dissertation provides support for the general physical-self concept scale which includes the self-concepts of muscle-tone, muscle-bulk, muscle-strength and body-fat. The distinction of muscle-tone relative to muscle-bulk and body-fatness requires future work to improve the conceptual clarity of these constructs.

Table of Contents

(continued)

(continued)

(continued)

(continued)

(continued)

(continued)

(continued)

(continued)

List of Tables

(continued)

(continued)

List of Figures

List of Symbols and Abbreviations

Abbreviation	Description
BF	Body-fat
BF1	Body-fat time-1
BF2	Body-fat time-2
BMI	Body mass index
BS	Bartlett test of sphericity
DMS	Drive for muscularity scale
EIS	Exercise identity scale
GLTEQ	Godin leisure time questionnaire
GPSCS	General physical self-concept scale
ICC	Intra-class coefficient
KMO	Kaiser-Meyer-Olkin
MANOVA	Multivariate Analysis of Variance
MDS	Multidimensional Scaling
MB	Muscle-bulk
MB1	Muscle-bulk time-1
MB2	Muscle-bulk time-2
METS	Metabolic equivalent of task
MS	Muscle-strength
MS1	Muscle-strength time-1
MS2	Muscle-strength time-2
MT	Muscle-tone

(continued)

Abbreviation	Description
MT1	Muscle-tone time-1
MT2	Muscle-tone time-2
PA	Principal axes
PSC	Physical self-concept

Development of a General Measure of Physical Self-Concept of Muscularity

The self has been studied for centuries by philosophers, sociologists, and psychologists. The complex nature of the self has resulted in multiple definitions and perspectives of what factors are found within the self and what factors affect and are affected by the self. The "self" is generally defined as the perceptions, thoughts, feelings, and beliefs one has in reference to him/herself (Leary & Tangney, 2003, p. 7). Furthermore, the self is "self-concept" which is defined as the beliefs or conceptualizations one has about oneself (Leary & Tangney, 2003, p. 7).

Self-concept is the aspect of the self that questions: "Who am I?" and "What am I like?" Self-esteem and self-concept are often used interchangeably but are conceptually different. Self-concept is generated from descriptive representations of the individual, whereas self-esteem is typically referred to as the beliefs and feelings of adequacy derived from an individual's evaluations of him/herself (Harter, 1999). Empirical evidence has not been able to distinguish between the self-description one has in a particular situation and the evaluation one has in this situation (Shavelson, Hubner, & Stanton, 1976). Shavelson et al.'s (1976) classical review of self-concept used the terms self-concept and self-esteem interchangeably, theorizing that self-concept is both descriptive and evaluative. This is likely the original source of the overlap between the two terms. While self-concept is not intended to be evaluative, many descriptions of the self have a positive or negative connotation according to social desirability. This paper will follow Shavelson's conception of self-concept being a descriptive and evaluative construct.

Self-concept is formed through experiences with the environment and influenced by the interactions and reinforcement of others which provides one with a general idea of his or her attributes (Bandura, 1978; Shavelson et al., 1976). Self-perceptions, behaviour and the environment have reciprocal influence on each other, thus self-concept is a product of the environment and behaviour as well as the environment and behaviour are influenced by self-concept (Bandura, 1978; Shavelson et al., 1976).

Self-concept is organized generally and according to specific domains that include social (Shavelson et al., 1976), physical (Marsh, Richards, Johnson, Roche, & Tremayne, 1994), and academic (Marsh, Byrne, & Shavelson, 1986) self-concepts, to name a few. Experiences within these different domains provides the individual with domain-specific information which is organized into more succinct self-perceptions describing one's attributes (Shavelson et al., 1976). Exposure to a variety of environments and individuals results in different domains of the self being used to guide one's role in a particular situation. Furthermore, different domains of the self become salient according to the environment and social contexts. For example, when people find themselves in an academic environment, their academic self-concepts may shape their interactions and behaviours rather than their artistic self-concepts. To accommodate the notion that different self-domains exist, the overall self-concept must be multidimensional. The multidimensional nature of self-concept suggests that each of these domains must be theorized, measured, and operationalized separately (Byrne & Shavelson, 1996). For example, if we look at the conceptual model proposed by Shavelson et al. (1976) (see Figure 1.1), the theorized multidimensional self-concept is operationalized such that different domains of the self can be defined and assessed as separate constructs that are further hypothesized to have independent influences on overall self-concept and potentially on other constructs.

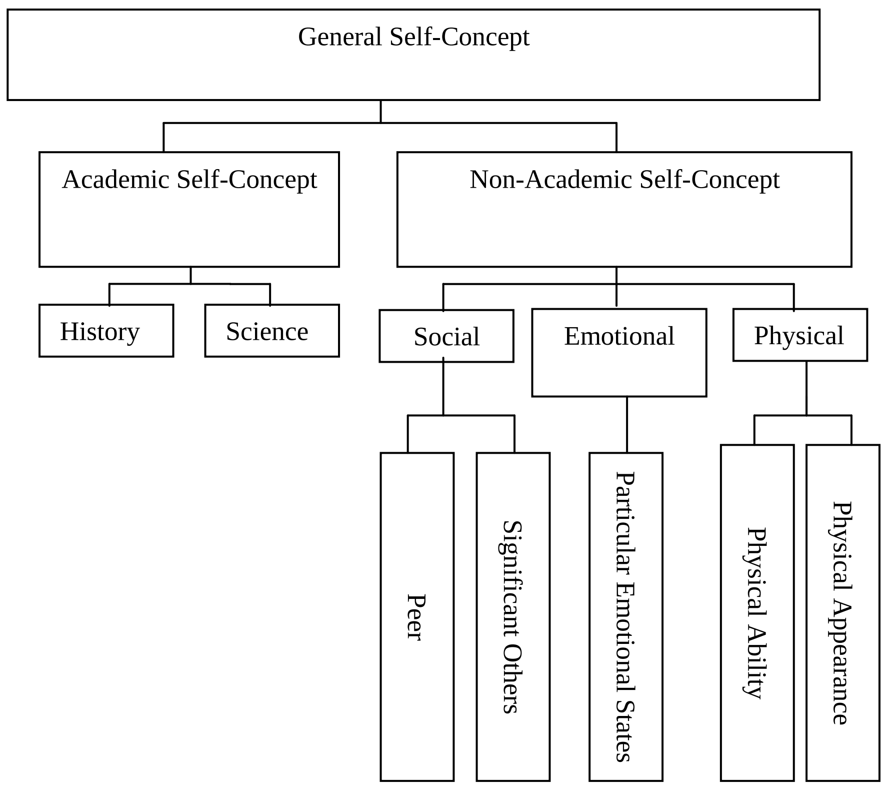

Figure 1.1. General self-concept. Extrapolated from Shavelson, R.J., Hubner, J.J., & Stanton, G.C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*, 407-441.

According to the Shavelson model, the structure of self-concept is hypothesized to be hierarchical. The hierarchical structure contains general self-concept at the apex or top of the model, which is split into two domains, the academic self-concept and the non-academic self-concept. The domain of academic self-concept further divides into more specific facets that include English, history, science, and math. The hierarchical nature of the Shavelson et al. (1976) self-concept model allows us to study different aspects of self-concept independently of each other in order to improve our understanding of the specific domains (Marsh & Yeung, 1998). More detailed hierarchical modelling of specific dimensions of self-concept has been conducted by different researchers (Byrne & Shavelson, 1996; Marsh, 1993a, 1996b; Vispoel, 1995). The Shavelson model laid the foundation for more recent theories and models that have added more specific domains of self-concept yet have retained the multidimensional, hierarchical structure of overall self-concept (Marsh, 1990, 1993a, 1996a). Some examples of domains that have been added include the physical self-concept (Marsh et al., 1994), the artistic self-concept (Vispoel, 1995), social self-concept (Bain & Bell, 2004), emotional self-concept (Marsh, Ellis, Parada, Richards, & Heubeck, 2005), and academic self-concept (Bong & Skaalvik, 2003). The more detailed modeling of the more specific self-concept domains and sub-domains, such as physical self-concept, have demonstrated good structural validity (Marsh, et al., 1994), criterion related validity evidence (Marsh, 1996b; Marsh & Redmayne, 1994) and the ability to distinguish between groups (Sonstroem, Speliotis, & Fava, 1992). For example, the physical self-concept sub-domain of physical competence was associated with physical activity participation and was significantly greater among exercisers than non-exercisers (Sonstroem et al., 1992). The focus of this paper will be the domain of the physical self-concept.

Physical Self-Concept: Theoretical Development

The physical self-concept has a unique position in the self-system due to the body's appearance, attributes, and abilities functioning as a substantial interface between a person and the world (Fox, 2000). The body plays a major role in social communication and may be used to express social status. For example, attractiveness of the body is positively related to increased confidence in social interactions (Eagly, Ashmore, Makhijani, & Longo, 1991; Nezlek, 1999); being perceived as more intelligent (Jackson, Hunter, & Hodge, 1995); being mentally healthy (Feingold, 1992); motivating others to befriend or date the attractive person (Peretti & Abplanalp, 2004); being more competent (Parks & Kennedy, 2007); and being perceived as earning more money (Frieze, Olson, & Russell, 1991). People are consciously or subconsciously aware of the social phenomena associated with physical appearance; therefore, having a positive self-concept regarding the body's physical appearance is an important aspect of socialization (Nezlek, 1999).

Perceptions of one's physical abilities impact participation in physical activities and have been demonstrated to distinguish between exercisers and non-exercisers (Sonstroem et al., 1992). A reciprocal effect exists between physical self-concept and exercise participation (Marsh, Papaioannou, & Theodorakis, 2006). Having a positive physical self-concept is related to increased levels of future physical activity participation, which supports the self-enhancement model

(Sonstroem, Harlow, & Salisbury, 1993). The self-enhancement model suggests that positive self-perceptions are associated with motivation to engage in particular behaviours. For example, feeling competent at exercise makes people want to demonstrate their competence in this area. As a result, this desire to demonstrate competence motivates them to exercise (Biddle & Mutrie, 2001). In addition, past physical activity participation is related to improvements in physical self-concept, which supports the skill development model (Marsh et al., 2006; Wilfley & Kunce, 1986). The skill development model suggests that physical self-concept is shaped by behavioural processes such as skill development (Crocker, Sabiston, Kowalski, McDonough, & Kowalski, 2006; Fox, 1997;) This reciprocal relationship means that physical activity and physical self-concept need to be developed together by exercise professionals to promote the long-lasting health benefits associated with ongoing exercise participation (Marsh et al., 2006). A positive physical self-concept has also been related to health benefits such as improvements in physical and mental well-being (Sonstroem & Potts, 1996) and reductions in behaviours that are associated with health risks (Rodriguez & Audrain-McGovern, 2005). In addition, physical self-concept has been shown to mediate the relationship between physical activity, sports participation, and self-esteem, which is inversely related to depression among adolescents (Dishman et al., 2006). Physical self-concept incorporates appearance-related and physical ability-related facets of the self that define specific aspects of one's physical self-conceptions and are associated with different forms of motivation, behaviour, and affect. The physical self-concept is an important social cognitive variable that can assist us in understanding the relationships between individuals, the environment, and their behaviour (Shavelson et al., 1976).

Physical self-concept is defined as the self-perception of one's appearance and physical capabilities (Marsh et al., 1994). Shavelson's model of self-concept includes physical self-concept as a non-academic domain and is, in turn, further separated into physical abilities and physical appearance (see Figure 1.2). Marsh (1996a) developed a more detailed multidimensional and hierarchical model of the physical self (see Figure 1.2). The multiple dimensions of the physical self-concept are theorized to include facets that assess self-perceptions of physical activity levels and competencies, appearance, physique, and health (Fox & Corbin, 1989; Marsh, 1996a). Marsh et al.'s (1994) multidimensional hierarchical model of physical self-concept is structured with global self-esteem at the apex; followed by general physical self-concept at the domain level; and strength, flexibility, endurance, physical activity, sport competence, coordination, health, body-fat, and appearance at the facet level (base of the model), shown in Figure 1.2 (Marsh, 1996). Self-esteem, which is at the apex of the model, represents overall positive feelings about the self in general. At the next level, the domain level, general physical self-concept represents feelings about the physical self. The nine facets (defined in Table 1.1) at the lowest level of the model are postulated to be less stable and more susceptible to environmental influences than the domain levels or the apex of the model (Sonstroem et al., 1992). For example, overall global physical self-concept may remain the same whether one is playing a sport in a competitive or recreational environment. In contrast, the perception of sport competence (at the facet level of the model) may fluctuate drastically depending on the type of sports environment in which one participates. The hierarchical structure of the model can assist researchers in understanding the volatility of the constructs at the facet level. The facets are more susceptible to fluctuate with changes in the environment and other external factors. This dissertation further explores the facet level located at the base of the physical self-concept hierarchy which is the most specific form of physical self-concept. Physical self-concept facets may be categorized as describing physical abilities or appearance aspects of the self.

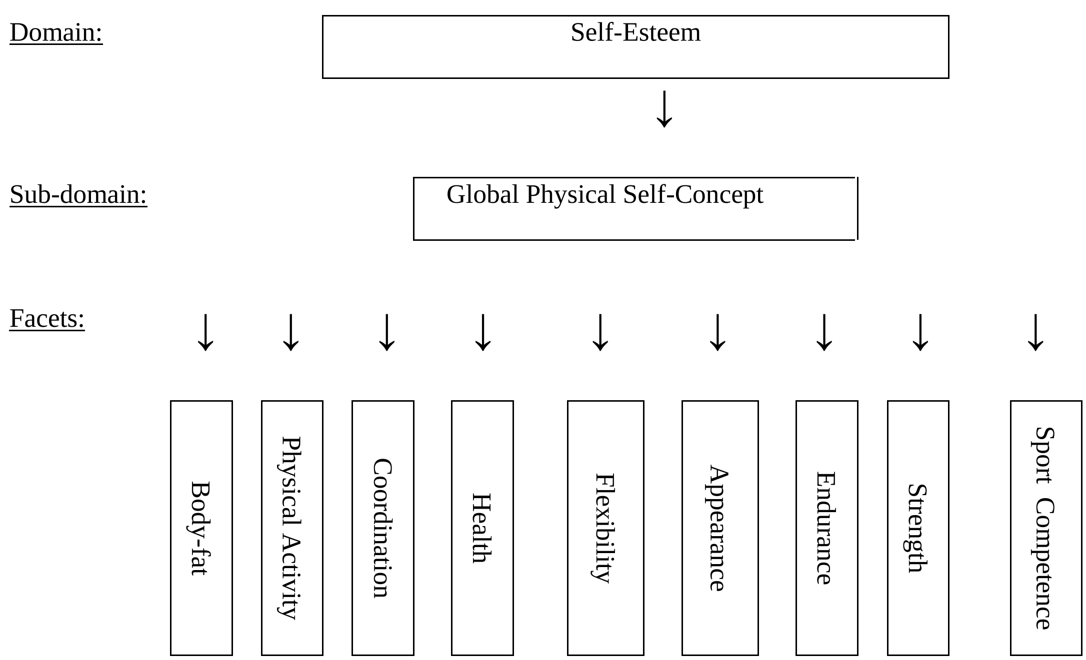

Figure 1.2. The Physical Self-Description Questionnaire (PSDQ; Marsh, 1996a). Extrapolated from Physical Self-Description Questionnaire: Psychometric properties and a multitrait-multimethod analysis of relations to existing

instruments, by Marsh, H.W., Richards, G.E., Johnson, S., Roche, L., & Tremayne, P., 1994, *Journal of Sport & Exercise Psychology*, *16*, 270-305.

Physical self-concept: Physical abilities. According to Marsh et al.'s (1994) theory, the facets associated with physical abilities include strength, flexibility, endurance, physical activity, sport competence, coordination, and health. The key components of physical abilities were originally identified and based on the physical fitness indicators from the Australian Health and Fitness Survey derived from Fleishman's (1964) structure of physical fitness (Marsh, 1996b). Different types of fitness tests have been used to provide convergent and discriminant validity evidence supporting the use of the physical ability facets (i.e., endurance, flexibility, strength, etc.) as a measure of the physical self-concept (Asci, 2005; Guerin, Marsh, & Famose, 2004; Marsh, 1996b; Tsorbatzoudis, 2005). Some of the fitness tests require skills that would be related to one or more perceptions of physical ability facets. For example, convergent validity evidence has been demonstrated as the perception of strength had a strong positive correlation with the pull-up, basketball throw and standing long jump scores among high school students (Marsh, 1996b). The strong positive relationships of all the physical ability facets with their corresponding fitness tests support the convergent construct validity and the lack of relationship between non-matching subscale and fitness tests provide discriminant validity evidence. The breadth of the physical ability facets within Marsh's framework is currently fairly comprehensive, six of the nine facets assess physical ability.

Table 1.1

Definitions of the Constructs Proposed by Marsh, Richards, Johnson, Roche & Tremayne, (1994) to Represent the Physical Self-Concept

Construct	Definition
Self-esteem	Assess overall positive feelings about the self in general.
Global physical self-concept	Assess positive feelings about the physical self.
Sport competence	Being good at sports, being athletic, having good sport skills.
Body-fat	Not being overweight, not being fat.
Physical activity	Being physically active, doing lots of physical activities regularly.
Coordination	Being good at coordinated movements, being able to do physical movements smoothly.
Health	Not getting sick often, getting well quickly when you are sick.
Flexibility	Being able to bend and turn your body easily in different directions.
Endurance	Being able to run a long way without stopping, not tiring easily when exercising hard.
Strength	Being strong, having a powerful body with lots of muscles.
Appearance	Being good-looking and having a nice face.

Note. Adapted from Physical Self-Description Questionnaire: Psychometric properties and a multitrait-multimethod analysis of relations to existing instruments, by Marsh, H.W., Richards, G.E., Johnson, S., Roche, L., & Tremayne, P., 1994, *Journal of Sport & Exercise Psychology*, *16*, 270-305.

Physical self-concept: Physical appearance. The facets of Marsh et al's (1994) theory of physical self-concept that are associated with physical appearance are appearance and body-fatness. The appearance facet includes self-assessments of facial features and perceptions of being good-looking. Body-fatness is the second facet Marsh uses to define the body's appearance. Body-fatness is one aspect of body composition related to physical fitness that is of interest to physical activity researchers.

Perceptions of body-fatness have predominantly been studied among women, although over the last two decades it has become a more prevalent topic among men. Negative perceptions of body-fatness among women are so rampant that it is considered a social norm for women of all ages to engage in "fat talk," which is the negative discussion of one's body (Nichter & Vuckovic, 1994). Fat talk can include engaging or joking with peers about one's "fat body" while trying to elicit validation from those peers (Britton, Martz, Bazzini, Curtin, & LeaShomb, 2006). Many studies have found that women perceive their body as fatter than the body they believe they ought to or ideally would like to have (Loitz, 2004; Szymanski & Cash, 1995). Although this negative self-concept of body-fatness is considered "normal" among women, these types of feelings can be associated with poor mental health states including depression and anxiety or eating disorders such as anorexia, binge eating, and bulimia. Much of the research assessing women's perceptions of body-fatness has derived from the study of psychological disorders such as body dysmorphia, eating disorders, and body image disorders (Thompson, Heinberg, Altabe, & Tantleff-Dubb, 2002).

Muscularity

Although the primary body dissatisfaction concerns among North Americans have traditionally been associated with a drive for thinness and fat loss, researchers have documented that men and women also experience dissatisfaction in regards to muscularity (Gray & Ginsberg, 2007). Despite this recognition, existing measures of physical self-concept (PSDQ; PSPP) do not include measures of perceived muscularity. Given this omission, it seems reasonable to propose the need to develop a scale that can assess muscularity as another component of physical appearance of the physical self-concept.

Historically, physical self-concept and body image research has focused on body-fatness, but since the 1990s, a growing body of research has focused on muscularity and hyper-morphology among men and women (e.g., McCreary & Sadava, 2001; McCreary & Sasse, 2000; Olivardia, Pope, Borowiecki, & Cohane, 2004; Pope et al., 2000). This research has focused on the drive for muscularity in men which has been compared to the drive for thinness found among women (Leit, Gray, & Pope, 2002; McCreary & Sasse, 2000; Pope, Olivardia, Borowiecki, & Cohane, 2001).

It has been suggested that the drive for muscularity is fuelled by images and attitudes perpetuated by the media and Western culture. Pope, Olivardia, Gruber, and Borowiecki (1999) documented trends among male action figure toys, such as G.I. Joe, from 1964 to 1994, and found that the muscularity depicted on the figures was unattainable for most men even with the use of ergogenic aids such as steroids and creatine. Changes in the ideal male physique were assessed in a study of centrefold models in *Playgirl* magazines (Leit, Gray, & Pope, & Gray, 2002; Spitzer, Henderson, & Zivian, 1999). Leit, Pope, Katz, and Oliva (2001) found that the male models have shed approximately 12 pounds of fat and added 27 pounds of muscle over the last 25 years. A third form of media study has found that men's bodies are being used more often in advertisements for products that are not related to the body than they were 50 years ago (Pope et al., 2001). According to Pope et al., men's bodies were used in 3% of advertisements in 1950, whereas men's bodies were being used in 35% of advertisements in 1990. The evidence gathered from these papers clearly suggests that a firm, muscular body is the desired ideal men's body displayed in Western media.

Studies have demonstrated that men tend to choose an ideal body that is slightly larger and more muscular than their current selves (Olivardia et al., 2004). Furthermore, men generally perceive that women prefer a more muscular and leaner body than women actually would choose for the ideal man (Olivardia et al., 2004). Within this same study, it was found that lower self-esteem was correlated to body dissatisfaction, muscle displeasure (i.e., the individual wants to be more muscular than he currently perceives himself to be), muscle belittlement (i.e., the individual thinks he is less muscular than he actually is), and feeling out of shape. Low self-esteem and low physical self-concept were more highly correlated with perceptions of muscularity than with body-fatness (Olivardia et al., 2004). The findings of these studies suggest that men (18 to 30 years) in Western culture perceive a greater external and internal pressure to meet the ideal of a hyper-muscular physique.

Muscularity research has predominantly examined male samples, although a small number of studies have assessed muscularity among women (Choi, 2003; Kyrejto, Mosewich, Kowalski, Mack & Crocker, 2008; Markula, 1995; Mosewich, Vangool, Kowalski, & McHugh, 2009). The primary research focus of these studies was to describe and explore body ideals among women and the exploration of the drive for muscularity. Qualitative studies by Markula (1995) and Mosewich et al. (2009) explored body ideals among women. Markula (1995) found that aerobicicers strived for toned, tight, thin and muscular bodies which was a hybrid of a feminine look and a strong, muscular appearance. Furthermore, much description of the desire to have “unbulky muscles”, and a toned body with long, sleek and firm muscles was documented. Mosewich et al. (2009) found that even female athletes were concerned about the accumulation of bulky, masculine muscles despite the athletic performance benefits.

North American studies have found that a drive for muscularity exists among some women (McCreary & Sasse, 2000; McCreary, Sasse, Saucier, & Dorsch, 2004). In a study with college aged females, 78% wanted to be more muscular, and 4% wanted to be less muscular (Jacobi & Cash, 1994). Kyrejto et al. (2008) found the drive for muscularity among women was similar to men’s drive for muscularity although women desired toned muscles rather than big, bulky muscles. Furthermore, correlates of the drive for muscularity, exercise and diet, were similar across gender (Kyrejto et al., 2005). Although the type of muscles idealized among men and women may differ, muscularity is an important aspect of the self-structure of the physical self-concept across gender.

The literature available provides a strong conceptual rationale to include muscularity as a facet of physical self-concept for men and women. Given that muscularity has been identified as a focal point of body ideals (Davis, Karvinen, & McCreary, 2005; Markula, 1995; Ridgeway & Tylka, 2005; Vartanian, Giant, & Passino, 2001), theories of physical self-concept should be updated to include muscularity. Consequently, measures that assesses muscularity according to levels of muscle-bulk and muscle-tone are required. To date, the PSPP (Fox and Corbin, 1989) and the PSDQ (Marsh, 1996a) are the two most commonly used measures of physical self-concept. Neither instrument includes a measure of muscularity. Developing a muscularity subscale(s) that may be used independently or with other physical self-concept instruments would improve the representativeness of the appearance aspects of physical self-concept research.

The purpose of this dissertation was to (a) develop a general measure of physical self-concept of muscularity among young adults, (b) assess the content relevance and representativeness of the items designed to measure the construct(s) of interest with expert panels, (c) produce internal structural validity evidence for the muscularity subscales, and (d) produce external criterion-related validity and reliability evidence for the muscularity subscales.

References

Asci, F. H. (2005). The construct validity of two physical self-concept measures: An example from Turkey. *Psychology of Sport and Exercise, 6*, 659-669. doi:10.1016/j.psychsport.2004.07.005

Bain, S. K., & Bell, S. M. (2004). Social self-concept, social-attributions, and peer relationships in fourth, fifth, and sixth graders who are gifted compared to high achievers. *Gifted Children Quarterly, 48*, 167-178.doi:10.1177/0016 98620404800302

Bandura, A. (1978). The self system in reciprocal determinism. *American Psychologist, 33*, 334-358. doi:10.1037/0003-066X.33.4.344

Biddle, S., & Mutrie, N. (2001). I can! Motivation through feelings of competence and confidence. *Psychology of physical activity: Determinants, well-being and interventions* (pp. 85-108). London, GB: Routledge.

Bong, M., & Skaalvik, E. M. (2003). "Academic self-concept and self-efficacy: How different are they really?". *Educational Psychology Review, 15*, 1-34. doi:10.1023/A:1021302408382

Britton, L. E., Martz, D. M., Bazzini, D. G., Curtin, L. A., & LeaShomb, A. (2006). Fat talk and self-orientation of body image: Is there a social norm for women to self-degrade? *Body Image, 3*, 247-254. doi:10.1016/j.body im.2006.05.006

Byrne, B. M., & Shavelson, R. J. (1996). On the structure of social self-concept for pre-, early, and late adolescents: A test of the Shavelson, Hubner and Stanton (1976) model. *Journal of Personality and Social Psychology, 70*, 599-613. doi:10.1037//0022-3514.70.3.599

Choi, P. Y. L. (2003). Muscle matters: Maintaining visible differences between women and men. *Sexualities, Evolution and Gender, 5*, 71-81. doi:10.1080/14616660310001632554

Crocker, P. R. E., Sabiston, C. M., Kowalski, K. C., McDonough, M. H., & Kowalski, N. (2006). Longitudinal assessment of the relationship between physical self-concept and health-related behavior and emotion in adolescent girls. *Journal of Applied Sport Psychology, 18*, 185-200. doi:10.1080/10413200600830257

Davis, C., Karvinen, K., & McCreary, D. R. (2005). Personality correlates of a drive for muscularity in young men. *Personality and Individual Differences, 39*, 349-359. doi:10.1016/j.paid.2005.01.013

Dishman, R. K., Hales, D. P., Pfeiffer, K. A., Felton, G., Saunders, R., Ward, D. S.,… Pate, R.R. (2006). Physical self-concept and self-esteem mediate cross-sectional relations of physical activity and sport participation with depression symptoms among adolescent girls. *Health Psychology, 25*, 396-407. doi:10.1037/0278-6133.25.3.396

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but …: A meta-analysis of research on the physical attractiveness stereotype. *Psychological Bulletin, 110*, 109-128. doi:10. 1037//0033-2909.110.1.109

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin, 111*, 304-341. doi:10.1037/0033-2909.111.2.304

Fox, K. R. (1997). The physical self and process in self-esteem development. In K. R. Fox (Ed.), *The physical self: From motivation to well-being* (pp. 111-139). Champaign, IL: Human Kinetics.

Fox, K. R., & Corbin, C. B. (1989). The physical self-perception profile: Development and preliminary validation. *Journal of Sport and Exercise Psychology, 11*(4), 408-430. Retrieved from http://journals.humankinetics. com/jsep

Frieze, I. H., Olson, J. E., & Russell, J. (1991). Attractiveness and income for men and women in management. *Journal of Applied Social Psychology, 21*, 1039-1057. doi:10.1111/j.1559-1816.1991.tb00458.x

Gray, J., & Ginsberg, R. (2007). An overview of psychological and cultural research and theory. In J. K. Thompson & G. Cafri (Eds.), *The muscular ideal: Psychological, social and medical perspectives* (pp. 14-42). doi:10. 1037/11581-001

Guerin, F., Marsh, H. W., & Famose, J. P. (2004). Generalizability of the PSDQ and its relationship to physical fitness: The European French connection. *Journal of Sport and Exercise Psychology, 26*(1), 19-38. Retrieved from http://journals.humankinetics.com/jsep

Harter, S. (1999). Contemporary issues and historical perspectives. *The construction of the self: A developmental perspective* (pp. 1-27). New York, NY: Guilford Press.

Jackson, L. A., Hunter, J. E., & Hodge, C. N. (1995). Physical attractiveness and intellectual competence: A meta-analytic review. *Social Psychology Quarterly, 58*, 108-122. doi:10.2307/2787149

Jacobi, L., & Cash, T. F. (1994). In pursuit of the perfect appearance: Discrepancies among self-ideal perceptions of multiple physical attributes. *Journal of Applied Social Psychology, 24*, 379-396. doi:10.1111/j.1559 -1816.1994.tb00588.x

Kyrejto, J. W., Mosewich, A. D., Kowalski, K. C., Mack, D. E., & Crocker, P. R. E. (2008). Men's and women's drive for muscularity: Gender differences and cognitive and behavioral correlates. *International Journal of Sport and Exercise Psychology, 6*, 69-84. doi:10.1080/1612197X.2008 .9671855

Leary, M. R., & Tangney, J. P. (2003). The self as an organized construct in the behavioural and social sciences. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 3-14). New York, NY: Guilford Press.

Leit, R. A., Gray, J. J., & Pope, H. G. J. (2002). The media's representation of the ideal male body: A cause for muscle dysmorphia? *International Journal of Eating Disorders, 31*, 334-338. doi:10.1002/eat.10019

Leit, R. A., Pope, H. G., Katz, D. L., & Oliva, P. (2001). Cultural expectations of muscularity in men: The evolution of Playgirl centerfolds. *International Journal of Eating Disorders*, *29*, 90-93. doi:10.1002/1098-108X(200101) 29:1<90::AID-EAT15>3.0.CO;2-F

Loitz, C. C. (2004). *The relationship among physical self-discrepancies, affect responses and behavioural intentions to exercise* (Unpublished master's thesis). University of Alberta, Edmonton, AB.

Markula, P. (1995). Firm but shapely, fit but sexy, strong but thin: The postmodern aerobicized female body. *Sociology of Sport Journal*, *12*, 424-453. Retrieved from http://journals.humankinetics.com/ssj

Marsh, H. W. (1990). The structure of the academic self-concept: TheMarsh/ Shavelson model. *Journal of Educational Psychology, 82*, 623-636. doi: 10.1037/0022-0663.82.4.623

Marsh, H. W. (1993a). Academic self-concept: Theory measurement and research. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 40, pp. 59-98). Hillsdale, NJ: Erlbaum.

Marsh, H. W. (1993b). The multidimensional structure of physical fitness: Invariance over gender and age. *Research Quarterly for Exercise and Sport, 64*, 256-274.

Marsh, H. W. (1996a). Physical Self-Description Questionnaire: Stability and discriminant validity. *Research Quarterly for Exercise and Sport*, *67*, 249-264.

Marsh, H. W. (1996b). Construct validity of Physical Self Description Questionnaire responses: Relations to external criteria. *Journal of Sport and Exercise Psychology, 18,* 111-131. Retrieved from http://journals. humankinetics.com/jsep

Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology, 80,* 366-380. Retrieved from http://www.apa.org/pubs/journals/edu/

Marsh, H. W., Ellis, A. L., Parada, R. H., Richards, G. E., & Heubeck, B. G. (2005). A short version of the self-description questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory analyses. *Psychological Assessment, 17*, 81-102. doi:10.1037/1040-3590.17.1.81

Marsh, H. W., & Famose, J. P. (2004). Generalizability of the PSDQ and its relationship to physical fitness: The European French connection. *Journal of Sport & Exercise Psychology, 26*, 19-38. Retrieved from http://journals.humankinetics.com/jsep

Marsh, H. W., Papaioannou, A., & Theodorakis, Y. (2006). Causal ordering of physical self-concept and exercise behaviour: Reciprocal effects model and the influence of physical education teachers. *Health Psychology*, *25*, 316-328. doi:10.1037/0278-6133.25.3.316

Marsh, H. W., & Redmayne, R. S. (1994). A multidimensional physical self-concept and its relations to multiple components of physical fitness. *Journal of Sport & Exercise Psychology, 16*, 43-55. Retrieved from http://journals.humankinetics.com/jsep

Marsh, H. W., Richards, G. E., Johnson, S., Roche, L., & Tremayne, P. (1994). Physical Self-Description Questionnaire: Psychometric properties and a multitrait-multimethod analysis of relations to existing instruments. *Journal of Sport & Exercise Psychology, 16*, 270-305. Retrieved from http://journals.humankinetics.com/jsep

Marsh, H. W., & Yeung, A. S. (1998). Top-down, bottom-up, and horizontal models: The direction of causality in multidimensional, hierarchical self concept models. *Journal of Personality, 75*, 509-527. doi:10.1037/0022-3514.75.2.509

McCreary, D. R., & Sadava, S. W. (2001). Gender differences in relationships among perceived attractiveness, life satisfaction, and health in adults as a function of body mass index and perceived weight. *Psychology of Men and Masculinity, 2*, 108-116. doi:10.1037//1524-9220.2.2.108

McCreary, D. R., & Sasse, D. K. (2000). An exploration of the drive for muscularity in adolescent boys and girls. *Journal of American College Health, 48,* 297-304. doi:10.1080/07448480009596271

McCreary, D. R., Sasse, D. K., Saucier, D. M., & Dorsch, K. D. (2004). Measuring the drive for muscularity: Factor validity of the drive for muscularity scale in men and women. *Psychology of Men and Masculinity, 5*, 49-58. doi:10.1037/1524-9220.5.1.49

Mosewich, A. D., Vangool, A. B., Kowalski, K. C., & McHugh, T. F. (2009). Exploring women track and field athletes' meanings of muscularity. *Journal of Applied Sport Psychology, 21*, 99-115. doi:10.1080/10413200802575742

Nezlek, J. B. (1999). Body image and day-to-day social interaction. *Journal of Personality, 67,* 793-817. doi:10.1111/1467-6494.00073

Nichter, M., & Vuckovic, N. (1994). Fat talk. In N. Sault (Ed.), *Many mirrors: Body image and social relations* (pp. 109-131). New Brunswick, NJ: Rutgers University Press.

Olivardia, R., Pope, H. G., Borowiecki, J. J., & Cohane, G. H. (2004). Biceps and body image: The relationship between muscularity and self-esteem, depression, and eating disorder symptoms. *Psychology of Men and Masculinity, 5*, 112-120. doi:10.1037/1524-9220.5.2.112

Parks, F. R., & Kennedy, J. H. (2007). The impact of race, physical attractiveness, and gender on education majors' and teachers' perceptions of student competence. *Journal of Black Studies, 37*, 936-943. doi:10.1016/S0041-1345(96)00567-2

Peretti, P. O., & Abplanalp, R. R. (2004). Chemistry in the college dating process: Structure and function. *Social Behavior and Personality, 32*, 147-154. doi: 10.2224/sbp.2004.32.2.147

Pope, H. G., Gruber, A. J., Mangweth, B., Bureau, B., deCol, C., Jouvent, R., & Hudson, J. I. (2000). Body image perception among men in three countries. *American Journal of Psychiatry, 157*, 1297-1301. doi:10.1176 /appi.ajp.157.8.1297

Pope, H. G., Olivardia, R., Borowiecki, J. J., & Cohane, G. H. (2001). The growing commercial value of the male body: A longitudinal survey of advertising in women's magazines. *Psychotherapy and Psychosomatics, 70*, 189-192. doi:10.1159/000056252

Pope, H. G., Olivardia, R., Gruber, A., & Borowiecki, J. (1999). Evolving ideals of male body image as seen through action toys. *International Journal of Eating Disorders, 26*, 65-72. doi:10.1002/(SICI)1098-108(1999907)26: 1<65::AID-EAT8>3.3.CO;2-4

Ridgeway, R. T., & Tylka, T. L. (2005). College men's perceptions of ideal body composition and shape. *Psychology of Men and Masculinity, 6*, 209-220. doi:10.1037/1524-9220.6.3.209

Rodriguez, D., & Audrain-McGovern, J. (2005). Physical activity, global physical self-concept, and adolescent smoking. *Annals of Behavioral Medicine, 30*, 251-259. doi:10.1207/s15324796abm3003_9

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*, 407-441. doi:10.2307/1170010

Sonstroem, R. J., Harlow, L. L., & Salisbury, K. S. (1993). Path analysis of self-esteem model across a competitive swim season. *Research Quarterly for Exercise and Sport, 64*(3), 335-342.

Sonstroem, R. J., & Potts, S. A. (1996). Life adjustment correlates of physical self-concept. *Medicine and Science in Sport and Exercise, 28*, 619-625. doi:10.1249/00005768-199605000-00014

Sonstroem, R. J., Speliotis, E. D., & Fava, J. L. (1992). Perceived physical competence in adults: An examination of the physical self-perception profile. *Journal of Sport and Exercise Psychology, 14*(2), 207-221. Retrieved from http://journals.humankinetics.com/jsep

Spitzer, B. L., Henderson, K. A., & Zivian, M. T. (1999). Gender differences in population versus media body size: A comparison over four decades. *Sex Roles, 40*, 545-565. doi:10.1023/A:1018836029738

Szymanski, M. L., & Cash, T. F. (1995). Body-image disturbances and self-discrepancy theory: Expansion of the Body-Image Ideals Questionnaire. *Journal of Social and Clinical Psychology, 14*, 134-146. doi:10.1521 /jscp.1995.14.2.134

Thompson, J. K., Heinberg, L. J., Altabe, M., & Tantleff-Dunn, S. (2002). *Exacting beauty: Theory, assessment and treatment of body image disturbance.* Washington, DC: American Psychological Association.

Tsorbatzoudis, H. (2005). Psychometric evaluation of the Greek Physical Self-Description Questionnaire. *Perceptual and Motor Skills, 101*, 79-89. doi: 10.2466/pms.102.3.703-718

Vartanian, L. R., Giant, C. L., & Passino, R. M. (2001). "Ally McBeal vs. Arnold Schwarzenegger": Comparing mass media, interpersonal feedback and gender as predictors of satisfaction with body thinness and muscularity. *Social Behaviour and Personality, 29*, 711-724. doi:10.2224/sbp.2001. 29.7.711

Vispoel, W. P. (1995). Self-concept in the arts: An extension of the Shavelson, Hubner and Stanton (1976) model. *Journal of Educational Psychology, 87*, 134-145. doi:10.1037/0022-0663.87.1.134

Wilfley, D., & Kunce, J. (1986). Differential physical and psychological effects of exercise. *Journal of Counseling Psychology, 33*, 337-342. doi:10.1037/00 22-0167.33.3.337

Study 1

Assessing the Content Validity of the Perception of Muscle-tone, Muscle-bulk, Muscle-strength and Body-fat Items Among an Expert Panel of Academics

Prior to the 1990s, body-ideals research traditionally focused on thinness and the absence of body-fat. However, research today shows that boys and men desire the muscular and "ripped" bodies found in popular magazines such as 'Men's Health', while women and girls prefer a defined or toned body as exemplified in 'Shape' and 'Self' magazines (Markula, 1995; Thompson & Cafri, 2007). A shift away from a purely anti-fat body towards a body ideal with some aspect of muscularity (e.g., muscle-bulk or muscle definition) requires an addition to the conventional measures of body image and physical self-concept (PSC). In order to study and assess muscle-bulk or muscle definition, the development of measures that assess muscular ideals are required. An instrument that is able to assess and interpret perceptions of the self-concept of muscularity independently and in conjunction with other facets of the physical self (e.g., body-fat, strength) would broaden our understanding of perceptions of muscularity and PSC. Therefore, the purpose of this study was to develop a set of items that would be used to assess overall PSC of muscle-tone and muscle-bulk.

Muscularity

Muscularity can be viewed from the perspective of muscle accumulation or the visibility of the muscle. The muscle accumulation perspective refers to the accumulation of muscle mass which can be viewed on a continuum of little muscle mass to a lot of muscle mass. Muscle-bulk is the term used to describe a large amount of muscle mass and does not take into consideration body-fat. In general, muscle-bulk is the accumulation of big, thick, and broad muscles. Men often idealize having large, broad shoulders and back, a thick neck, and a big chest and biceps (Ridgeway & Tylka, 2005; Thompson & Cafri, 2007). Men generally focus on the muscle-bulk of the upper body but some strive for bigger upper legs and calves (Ridgeway & Tylka, 2005). In contrast, women generally fear possessing bulky muscles (Gruber, 2007) and research suggests that female aerobic exercisers actually want "unbulky" muscles (Markula, 1995). Mosewich, Vangool, Kowalski and McHugh (2008) found that female track athletes (sprinters, hurdlers, horizontal jumper, long distance runners and a pole vaulter) also wanted to limit the amount of muscle. It appears that the amount of muscle mass idealized varies according to gender. Despite women not wanting to have big, bulky muscles, women do desire to be more muscular (Jacobi & Cash, 1994)

The second aspect of muscularity is the visibility of the muscles. Muscle-tone refers to the visibility of the muscle and describes having defined, "cut", "chiseled" or sculpted muscles. Muscle-tone takes into consideration the development of muscle as well as the amount of body-fat. In order for a muscle to be defined or "cut", muscle needs to be present and fat needs to be at a minimum level. A toned body among men generally refers to having a moderate or higher level of muscle development with very little body-fat to view the cut, sculpted or chiseled "six-pack" or the defined arms and chest (Ridgeway & Tylka, 2005).

Women also desire muscle-tone, although they prefer their muscles to appear long, lean and defined (Markula, 1995; Mosewich et al., 2008). Consequently, women appear to desire toned muscles with some muscle development and little body-fat.

Instrument Development

The construction of the operational definitions is a critical step in instrument development as the operational definition gives a precise indication of the characteristics and properties of a concept, as well as indicators of how to identify these characteristics (Bless & Higson-Smith, 2000). Within the literature, muscularity has generally possessed a singular meaning which has been problematic, as one term cannot capture the complexities of male and female body ideals (Blood, 2005; Kyrejto, Mosewich, Kowalski, Mack & Crocker, 2008). The lack of distinction in the literature between muscle-tone and muscle-bulk has been problematic. A review of the muscularity literature (Butler & Ryckman, 1993; Dworkin, 2003; Evans, Cotter, & Roy, 2005; Frew & McGillivray, 2005; Haravon, 2002; Lenart, Goldberg, Bailey, Dallal, & Koff, 1995; Markula, 1995; Marsh, 1996a; Marsh, Richards, Johnson, Roche, & Tremayne, 1994; Ridgeway and Tylka, 2005) examining the terms used by researchers and study participants to describe muscularity was conducted and utilized to develop the operational definitions of muscle-tone and muscle-bulk.

The literature used to develop the muscularity operational definitions included qualitative and quantitative studies among men and women (Butler & Ryckman, 1993; Dworkin, 2003; Evans, Cotter, & Roy, 2005; Frew &

McGillivray, 2005; Haravon, 2002; Lenart et al., 1995; Markula, 1995; Marsh, 1996a; Marsh, et al., 1994; Ridgeway & Tylka, 2005). The constructs of muscle-strength and body-fat have been previously described and used by other researchers, therefore these operational definitions were examined and incorporated into the current operational definitions (see Fox, 1990; Fox & Corbin, 1989; Marsh et al., 1994). The definitions take into consideration the language and terms used by other muscularity and PSC researchers to define muscle-tone, muscle-bulk, muscle-strength and body-fat. The operational definitions developed for this study were as follows:

Muscle-tone: The self-perception of the degree to which muscles look lean, firm and defined.

Muscle-bulk: The self-perception of the size and mass of muscles.

Muscle-strength: The self-perception of the presence or absence of strength and the ability to lift heavy objects.

Body-fat: The self-perception relating to the amount of body-fat.

The operational definitions were used to guide item content for the constructs of muscle-tone, muscle-bulk, muscle-strength and body-fat. Eight items were initially developed to measure each construct. The items were developed to represent muscle-tone and muscle-bulk according to young adults (18 to 29 years of age). Although eight items were developed to measure each construct, the goal at the end of this dissertation was to have a minimum of four items per subscale to measure each of the constructs. See Table 2.1 for the item pool for the General

Physical Self-Concept Scale (GPSCS) that was designed to assess muscle-tone, muscle-bulk, muscle-strength and body-fat.

Table 2.1

Items Proposed to Represent the Subscales of Muscle-tone, Muscle-bulk, Muscle-strength and Body-fat in the General Physical Self-Concept Scale

Subscale	Item
Muscle-tone	1. My muscles are toned.
	2. My body is firm.
	3. My muscles appear lean.
	4. My muscles are well defined.
	5. My body looks muscular.
	6. I have a large amount of muscle mass on my body.
	7. My body looks flabby.
	8. My body is soft.
Muscle-bulk	9. I have large muscles.
	10. I have bulky muscles.
	11. I have huge muscles.
	12. I have small muscles.
	13. I have scrawny muscles.
	14. I have no muscles.
	15. My muscles are slender.
	16. I have bulging muscles.

(continued)

Subscale	Item
Muscle-strength	17. I am physically strong.
	18. I am physically weak.
	19. I can lift heavy objects.
	20. I cannot lift heavy objects.
	21. I am physically powerful.
	22. I am physically feeble.
	23. My muscles are strong.
	24. My muscles are weak.
Body-fat	25. My body is fat.
	26. My body is skinny.
	27. My body is large.
	28. My body is thin.
	29. I have a plus-sized body.
	30. I have a lean body.
	31. I am chubby.
	32. I have a plump body.

The focus of Study 1 was to establish content validity evidence. One method of establishing content validity evidence is to have an expert panel judge the extent to which the content of each item is pertinent to the target construct it was intended to measure (DeVellis, 1991; Dunn, Bouffard, & Rogers, 1999; Gotwals & Dunn, 2009; Haynes, Richards, & Kubany, 1995).

Content relevance. Content *relevance* refers to “the degree to which the content (or subject matter) contained within a test item is representative of the targeted construct that the item is designed to measure” (Dunn et al., 1999, p.16). Although there are multiple methods of assessing item content relevance (see Crocker, 1997, Hambleton, 1978; Messick, 1995; Wynd, Schmidt & Schaefer, 2003), systematic approaches that assess content relevance are often overlooked in the scale construction process (Dunn et al., 1999; Hambleton, 1980; Messick, 1989). In this study, an expert panel of academics judged the ‘degree of match’ between the newly constructed items and the constructs they were designed to measure (see Gotwals, 2006; Hellsten, 2005). Content relevance of an instrument should be established prior to use of the instrument for assessment purposes (Hambleton, 1980).

Content representativeness. A second aspect of content validity is content *representativeness*. Content representativeness refers to the degree to which an item pool adequately covers the breadth of the identified aspects of a construct (Crocker & Algina, 1986). According to Messick (1989), to achieve representativeness or “domain coverage,” the boundaries of the domain need to be taken into consideration. Therefore, there were two components to account for in content representativeness:

1) the individual items being within the domain boundaries of the construct they represent and

2) the set of items representing a construct covers the breadth of the construct.

The constructs of interest in this study were muscle-tone, muscle-bulk, muscle-strength and body-fat which were considered narrow constructs relative to constructs such as overall PSC.

Method

Participant Characteristics

Academics with psychometric and PSC expertise were approached to be members of the expert panel. Eighteen academics (males n = 11; female n = 7) with a PhD participated in the study. Participants included 2 assistant professors, 6 associate professors, and 10 professors.

Sampling Procedures

The composition of the expert panel followed the guidelines and recommendations previously developed for scale construction (see Crocker & Algina, 1986; Dunn et al., 1999; Fitzpatrick, 1983). The academic experts were recruited due to their knowledge and expertise in psychological theory and scale development. The expert panel that judged the new items consisted of sport and exercise psychologists with a Ph.D. who had either published or presented in peer review settings on topics related to muscularity, physical self-concept, and psychometrics.

The researcher completed a search for the terms physical self-concept, physical self-perception, body image and muscularity in PSYCH INFO from 2000 to May, 2010. Forty-six potential expert panel members were identified and the researcher and supervisor agreed that forty of the forty-six would be appropriate to include as experts. Five potential experts were deemed to be inappropriate due to possible language issues and one had a conflict of interest in this study.

Forty experts were initially contacted by electronic mail and invited to participate in the study (see Appendix A for a copy of the recruitment email). [1] Experts were sent an email that contained a very brief description of the study, identified why they were selected for this study, informed them that the survey should take approximately 15 minutes, invited them to participate in the study by selecting the hyperlink that was embedded in the email and informed them that the survey needed to be completed within three weeks from the time the original email was sent. This hyperlink directed the experts to the information letter for the study followed by the survey package (see Appendix B for a copy of the information letter and Appendix C for a copy of the online survey package). The participant's consent was implied by the overt action of opening and completing the survey. Prior to starting this study, ethical approval was provided by an institutional research ethics board.

[1] Of the forty experts who were contacted to participate in this study two emails were returned due to the expert leaving their academic position (1= retired and 1= took a clinical position), three out of office replies due to travel or maternity leave, and one email was sent to the researcher stating that he was unable to complete this survey at this time. Eighteen experts (45%) completed the survey, six non-completers (15%) sent a message to the researcher, and sixteen participants (40%) did not complete the survey and did not contact the researcher. The response rate was lower than expected. However, according to Lynn (1986) and Aiken (1985) our sample size was adequate for the planned analyses. Lynn (1986) recommends a sample size ranging from five to ten experts for content validity studies, and Aiken (1985) suggests that 25 experts would be considered a large sample. Thus, 18 was deemed to be acceptable for the purposes of this study.[1]

The experts were asked to respond to demographic questions, content relevance and representativeness questions. Once three weeks from the delivery of the email had elapsed, the survey site was closed to further responses and the data were analyzed.

Measures

Demographics. The first part of the survey gathered demographic information from the expert panel which included sex, highest degree of education, and current academic position. The experts were asked to describe their current research area, indicate whether they have used PSC inventories and respond to a series of questions regarding their publication record in PSC research. See Appendix C, Part 1 for a copy of the demographic questions.

Content relevance of the general physical self-concept scale (GPSCS). The second section of the survey included questions pertaining to the content relevance of the 32 GPSCS items according to the constructs they were proposed to represent. The experts were asked to read the operational definition of the construct of interest (e.g., muscle-tone), then judge the degree of match between each item and the construct it was designed to measure. Ratings were made on a 5-point scale ranging from a poor match (1), fair match (2), good match (3), very good match (4), to excellent match (5). A space was given after each judgment to record any comments the expert had about that specific item. The judges were presented with eight items for each of the four constructs (i.e., muscle-tone, muscle-bulk, body-fat, and muscle-strength). This resulted in thirty-two items being judged within this section. See Appendix C, Part 2 for the complete GPSCS content relevance survey.

Content representativeness of the general physical self-concept scale (GPSCS). The third section of the survey included questions pertaining to content representativeness. The experts were asked how well the group of eight items assessed all identifiable aspects of the construct they were intended to measure. The survey provided the experts with the operational definition of the construct followed by the pool of items. The experts were asked: (a) the degree to which the item pool adequately covered the construct of interest, with the responses ranging from poor representativeness (1), fair representativeness (2), good representativeness (3), very good representativeness (4), to excellent representativeness (5); (b) if any additional items should be included; and (c) if any of the items were redundant or could be eliminated. The experts were provided with text boxes to comment on the content representativeness of each item set. See Appendix C, Part 3 for the complete GPSCS content representativeness survey.

Statistical Methods

Data were entered into SPSS 18 and analyzed. The demographic information was analyzed to confirm the panel members qualified as academic experts. Next the data set was screened for missing data; participants responding to a minimum of 90% of the items were included in the analyses.

Screening for discrepant raters. Given the small number of judges that are typically used in content relevance and representativeness studies, aberrant judges can adversely influence the central tendency scores and so must be removed (Hambleton, 1984). The experts' responses were examined to identify potentially aberrant judges by looking at each judge's response discrepancy from the median (JDM) (Hellsten, 2005). The JDM was measured to assess the degree of disagreement between each of the judge's responses and the corresponding median response (Hellsten, 2005). If the judges were in perfect agreement the JDM score would be zero. Conversely, the greater the disagreement among judges the larger the JDM score. The JDM calculation is as follows:

$$JDM_j = \sum_{k-1}^{k} |X_{kj} - Md_k|$$

where X_{kj} is the rating given by judge j to item k;

Md_k is the median of the rating given by the *J* judges to item *k*;

K is the number of items; and

$|X_{kj} - Md_k|$ is the absolute value between the rating given by judge *j* to item *k* and the median of the ratings given by the *J* judges to item *k* (Rogers, 2010).

Judges who deviated from others to a degree that undermined the validity of the evaluation were considered aberrant and removed from the quantitative analysis (Hambleton, 1984). Potential discrepant raters are initially identified as judges who have aberrantly high JDM scores and who differ substantially from the majority of judges (Gotwals, 2006; Hellsten, 2005; Rogers, 2010). A specific cut-off point for an aberrant judge is normally determined at the time of the JDM computation rather than a priori (Gotwals, 2006; Hellsten, 2005; Rogers, 2010).

Item fit analysis (item relevance). Item fit (or item relevance) was assessed according to: (a) the central tendency of the judges' ratings for the item, and (b) the item-content validity coefficient (Rogers, 2010).

a) Item fit. Two measures of central tendency, the mean and the median, were examined to assess the fit of each item to the construct they were intended to measure. The mean is the most commonly used measure of central tendency. One limitation of using the mean as a measure of central tendency is that the mean is overly sensitive to outliers in small samples; therefore under these circumstances the median is preferred. A score ≥ 2.70 *SD's* from the mean was considered an outlier (Glass &Hopkins, 1996, p.27). No outliers were identified upon inspection of the item fit. Both mean and median are reported to assess the fit of the items to the corresponding construct. However, according to Glass and Hopkins (1996) the mean is the preferred indicator of central tendency given the characteristics of these data. A higher central tendency score suggests a better fit according to the judges' ratings. According to the response scale, a score of ≥ 3 or greater was considered a good fit or better.

b) *Item-content validity coefficient.* Aiken's (1985) item content validity coefficient (*V*) is a statistical test that analyzes the relevance judgment scores to assess the degree of validity for each item as a representative of the corresponding construct. More specifically, Aiken's *V* assesses the degree to which the judges, as a group, indicate that each item measures the targeted domain (Dunn et al., 1999). The relevance ratings of a single item to the construct of interest by *n* experts are examined (Aiken, 1985). The rating scale for this study included five successive integers that ranged from one to five. The formula for Aiken's *V* is:

$$V = \frac{S}{n_r(c-1)},$$

where n = the number of judges,
c = the number of rating categories,
$S = \sum(r_{jk} - lo)$, lo = the lowest possible validity category,
r = judges j content validity rating item k.

Aiken's *V* (V-coefficient) scores range from 0 to 1, with a higher value indicating that the item had higher content validity. A higher content validity score suggests that the item is more relevant to the construct. Statistical significance of Aiken's *V* is associated with the number of rating categories and the number of experts. A right-tailed probability table was provided by Aiken (1985, p.134) in order to determine the statistical significance of the Aiken's *V* score. Items that demonstrated high central tendencies and a significant *V*-coefficient are deemed to have good fit to the construct or good content relevance.

Item representativeness analysis. The median and mean scores for item representativeness were examined. Higher mean and median scores on the item representativeness scores were deemed good and considered for retention. The qualitative feedback was used to make improvements to the items.

Results

Data Analysis

Content relevance analysis of the general physical self-concept scale.

Screening for discrepant raters. The prescreening of judges' ratings of the GPSCS identified two aberrant judges. The first judge was removed due to a poor response rate on the items. This respondent completed 62.5% of the quantitative questions which is below the minimum criterion of 90% completion (20 of the 32 quantitative questions were completed). A second aberrant judge was identified upon inspection of the JDM scores due to an extreme disagreement with the group of judges (JDM scores were 2.58 standard deviations from the mean and 2.74 standard deviations from the median). The aberrant judges' qualitative feedback was retained for the qualitative analysis. The content relevance scores for the 17 judges are reported in Table 2.2; (item means, item standard deviations, item medians, and judge's JDM scores). The content relevance scores of the 16 judges that were used for further analysis are reported in Table 2.3.

Quantitative assessment of content relevance ratings.

a) Item fit (item relevance):

Muscle-tone: The scores for the muscle-tone items identified three items that were considered to be a fair match (*Mdn*s = 2, *M*s = 2.13 to 2.63), three items that were a good match according to the mean scores (*M*s = 3.00 to 3.67) and two items were a good match according to the medians (*Mdn*s = 3), and two items that were considered to be a very good match according to the means (*M*s = 4.06 to 4.13) and three items according to the medians (*Mdn*s = *4*).

Muscle-bulk: The median and mean scores for the muscle-bulk items identified two items that were considered to be a fair match (*Mdn*s = 2, *M*s = 2.25 to 2.31), one item that was a good match according to the median scores (*Mdn* = 3) and five items according to the mean scores (*M*s = 3.12 to 3.94), and four items that were considered to be a very good match according to the median (*Mdn*s = 4) and one item according to the mean score (*M*s = 4.31).

Muscle-strength: The median scores for the muscle-strength items identified two items that were a good match (*Mdn*s = 3) and mean scores identified five items that were a good match (*M*s = 3.18 to 3.88), three items that were considered to be a very good match (*Mdn*s = 4, *M*s = 4.25 to 4.75) and three items that were considered to be an excellent match according to the median scores (*Mdn*s = 5).

Body-fat: The mean scores for the body-fat items identified one item that was a poor score (*M* = 1.67). The median and mean scores for the body-fat items identified one item that was a fair match (*Mdn* = 2, *M*s = 2.67). Two items were considered to be a good match according to the median scores (*Mdn*s = 3.0 and 3.5) and five items according to the mean scores (*M*s = 3.25 to 3.88). Five items were considered to be a very good match according to the median scores (*Mdn*s = 4.0) and one according to the mean scores (*M* = 4.19).

When examining the median scores for the items, 26 items were considered good to excellent, while six items were considered fair. According to the mean scores for the items, 25 items were considered a good to excellent match, while seven items were considered poor to fair match. These analyses provided the initial identification of items (1 poor and 6 fair items) that might need modification or may need to be removed from the list of items representing the construct of interest (items 6, 7, 8, 14, 15, 27 and 29).

Table 2.2

Summary of Seventeen Judges' Content Relevance Ratings

		Expert Panel Judge Identification Number																			
Constructs	Items	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	*M*	*SD*	*Mdn*
Tone	1	5	4	4	3	4	4	4	5	5	3	5	4	3	5	4	3	4	4.06	.75	4
	2	5	3	4	3	3	2	3	2	4	3	2	4	3	4	2	2	5	3.17	1.02	3
	3	5	3	4	5	3	4	2	3	3	3	2		4	5	4	4	5	3.69	1.01	4
	4	5	4	4	3	4	4	2	5	4	3	5	5	4	5	4	4	5	4.12	.86	4
	5	2	1	4	2	4	3	2	4	3	3	5	2	4	5	2	2	4	3.06	1.20	3
	6	1	1	4	1	3	4	2	2	2	2	1	1	3	4	2	2	3	2.24	1.02	2
	7	1	2	1	2	3	2	4	4	3	1	4	2	2	4	2	3	3	2.53	1.07	2
	8	1	2	1	2	2	2	3	3	3	1	3	2	2	3	2	1	3	2.12	.78	2
Bulk	9	5	5	3	4	4	4	4	4	4	4	4	4	4	5	4	5	5	4.24	.56	4
	10	3	4	1	4	4	4	4	5	4	3	5	4	3	4	4	4	2	3.65	1.00	4
	11	4	4		4	4	3	3	5	4	3	5	3	4	5	4	3	5	3.94	.77	4
	12	4	4	1	4	4	3	3	4	3	2	5	3	4	5	4		4	3.56	1.03	4
	13	3	3	1	4	2	2	3	3	3	2	5	3	4	4	3	4	2	3.00	1.00	3
	14	2	2	1	3	1	2	1	4	3	1	5	2	1	2	4	2	2	2.24	1.20	2
	15	4	2		2	4	3	2	2	2	1	2	1	1	4	2	2	2	2.25	1.00	2
	16		2	2	2	4	4	4	3	3	4	5	2	4	5	4	4	5	3.56	1.09	4
Strength	17	5	5	4	5	5	4	5	5	5	4	5	4	4	5	5	5	5	4.71	.47	5
	18	4	5	1	5	5	2	5	5	3	3	5	4	3	5	5	5	4	4.06	1.25	5
	19		4	4	2	3	4	5	5	3	1	5	4	4	5	4	4	5	3.88	1.15	4
	20	5	4	1	2	3	2	4	5	3	1	5	1	3	5	4	4	3	3.24	1.44	3
	21	5	4	2	3	4	4	5	4	2	3	4	3	4	2	5	2	4	3.53	1.07	4
	22	3	4	1	3	4	2	5	4	3	3	2	2	3	2	5	2	4	3.06	1.14	3
	23	5	5	2	5	4	2	5	5	4	4	5	4	4	5	5	3	5	4.24	1.03	5
	24	4	5	1	5	4	2	5	5	3	2	5	3	3	5	5	3	3	3.71	1.31	4

(continued)

		Expert Panel Judge Identification Number																			
Constructs	Items	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	*M*	*SD*	*Mdn*
Fat	25	5	5	4	5	4	4	5	5	4	4	1	4	4	5	4	3	5	4.18	1.02	4
	26	4	2	1	5	4	4	4	5	3	4	1	4	3	5	3	4	3	3.47	1.23	4
	27	1	1	2	2	1	2	1	2	1	1	2	2	2	2		2	3	1.69	.60	2
	28	4	2	1	5	4	2	4	5	2	2	1	3	3	3	4	4	4	3.12	1.27	3
	29	4	1	3	2	2	2	1	4	3	3	1	3		4	5	2	3	2.69	1.20	3
	30	4	4	3	2	4	4	3	4	2	2	5	4	3	5	4	4	4	3.59	.94	4
	31	5	3	3	5	4	4	3	4	4	3	3	4	4	5	4	3	4	3.82	.73	4
	32	5	3	1	4	4	4	2	4	4	4	4	4	4	5	4	3	3	3.65	1.00	4
JDM		23	19	45	22	14	20	26	23	20	29	37	17	14	31	13	20	23			

Note. JDM = Judge's discrepancy from the median. The mean JDM score was 23.29 (*SD* = 8.40) and the median JDM score was 22. See Table 2.1 for a list of the items.

Table 2.3

Summary of Sixteen Judges' Ratings

		Judges																			
Constructs	Items	1	2	4	5	6	7	8	9	10	11	12	13	14	15	16	17	*M*	*SD*	*Mdn*	Aiken's V
Tone	1	5	4	3	4	4	4	5	5	3	5	4	3	5	4	3	4	4.06	.77	4	0.63
	2	5	3	3	3	2	3	2	4	3	2	4	3	4	2	2	5	3.13	1.03	3	0.41
	3	5	3	5	3	4	2	3	3	3	2		4	5	4	4	5	3.67	1.05	4	0.48
	4	5	4	3	4	4	2	5	4	3	5	5	4	5	4	4	5	4.13	.89	4	0.64
	5	2	1	2	4	3	2	4	3	3	5	2	4	5	2	2	4	3.00	1.21	3	0.45
	6	1	1	1	3	4	2	2	2	2	1	1	3	4	2	2	3	2.13	1.03	2	0.25
	7	1	2	2	3	2	4	4	3	1	4	2	2	4	2	3	3	2.63	1.03	2	0.36
	8	1	2	2	2	2	3	3	3	1	3	2	2	3	2	1	3	2.19	.75	2	0.25
Bulk	9	5	5	4	4	4	4	4	4	4	4	4	4	5	4	5	5	4.31	.48	4	0.67*
	10	3	4	4	4	4	4	5	4	3	5	4	3	4	4	4	2	3.81	.75	4	0.59
	11	4	4	4	4	3	3	5	4	3	5	3	4	5	4	3	5	3.94	.77	4	0.61
	12	4	4	4	4	3	3	4	3	2	5	3	4	5	4		4	3.73	.80	4	0.50
	13	3	3	4	2	2	3	3	3	2	5	3	4	4	3	4	2	3.12	.89	3	0.44
	14	2	2	3	1	2	1	4	3	1	5	2	1	2	4	2	2	2.31	1.20	2	0.27
	15	4	2	2	4	3	2	2	2	1	2	1	1	4	2	2	2	2.25	1.00	2	0.22
	16		2	2	4	4	4	3	3	4	5	2	4	5	4	4	5	3.67	1.05	4	0.58
Strength	17	5	5	5	5	4	5	5	5	4	5	4	4	5	5	5	5	4.75	.45	5	0.78**
	18	4	5	5	5	2	5	5	3	3	5	4	3	5	5	5	4	4.25	1.00	5	0.67*
	19		4	2	3	4	5	5	3	1	5	4	4	5	4	4	5	3.87	1.19	4	0.59
	20	5	4	2	3	2	4	5	3	1	5	1	3	5	4	4	3	3.37	1.36	3	0.45
	21	5	4	3	4	4	5	4	2	3	4	3	4	2	5	2	4	3.63	1.03	4	0.52
	22	3	4	3	4	2	5	4	3	3	2	2	3	2	5	2	4	3.19	1.05	3	0.44
	23	5	5	5	4	2	5	5	4	4	5	4	4	5	5	3	5	4.38	.89	5	0.69*
	24	4	5	5	4	2	5	5	3	2	5	3	3	5	5	3	3	3.88	1.15	4	0.58

(continued)

Constructs	Items	Judges 1	2	4	5	6	7	8	9	10	11	12	13	14	15	16	17	*M*	*SD*	*Mdn*	V
Fat	25	5	5	5	4	4	5	5	4	4	1	4	4	5	4	3	5	4.19	1.05	4	0.64
	26	4	2	5	4	4	4	5	3	4	1	4	3	5	3	4	3	3.63	1.09	4	0.56
	27	1	1	2	1	2	1	2	1	1	2	2	2	2		2	3	1.67	.62	2	0.11
	28	4	2	5	4	2	4	5	2	2	1	3	3	3	4	4	4	3.25	1.18	3.5	0.47
	29	4	1	2	2	2	1	4	3	3	1	3		4	5	2	3	2.67	1.23	3	0.30
	30	4	4	2	4	4	3	4	2	2	5	4	3	5	4	4	4	3.62	.96	4	0.53
	31	5	3	5	4	4	3	4	4	3	3	4	4	5	4	3	4	3.88	.72	4	0.59
	32	5	3	4	4	4	2	4	4	4	4	4	4	5	4	3	3	3.81	.75	4	0.58
JDM		22.5	19.5	21.5	13.5	20.5	25.5	22.5	20.5	29.5	37.5	17.5	14.5	31.5	12.5	19.5	22.5				

Note. JDM = Judges discrepancy from the median. See Table 2.1 for a list of the items.
* $p < .05$. ** $p < .01$.

b) Item-content validity coefficient. For 16 raters and five rating categories, a V-coefficient of .72 corresponds to a *p* value of .008, while a V-coefficient of .66 corresponds to a *p* value of .046. Therefore, the V-coefficient of .66 as the cutoff value was used. According to the V-coefficient scores, only four of the 32 items were significant. These results indicate that for all but four items (9, 17, 18, 23), the ratings provided by the judges for an item on its proposed domain are not significantly higher than ratings that could occur due to chance. Although, the V-coefficient scores did not provide strong statistical support for the item fit, the item-relevance scores in the previous analysis identified most of the items as having mean and median scores of 'good match' to 'very good match'. These results demonstrate that the GPSCS items have potential to assess their matching construct but will require some modifications to items and removal of other items to improve the content reference of the items.

Qualitative Feedback of Content Relevance Ratings. Qualitative feedback for the 32 items is presented in Table 2.4. The first concern that appears to have been raised in the qualitative feedback relates to the precision of item wording to reflect the operational definitions. More specifically, the operational definitions of muscle-tone, muscle-bulk and body-fat conceptualize the constructs as capturing the "appearance" of muscles or the body. For example, an item describing the appearance of the body includes: "My muscles appear lean" (Item 3). Other items within the instrument refer to how the body or muscle "is", for example "My muscles are toned" (Item 1). These two different methods of phrasing the item may have different meaning to respondents, therefore all body-fat, muscle-tone and muscle-bulk items were re-phrased to take into account the appearance of the muscle or body. These recommendations are consistent with Shavelson's (1976) conceptualization of self-concept being a self-judgment of a particular dimension of the self. These items provide a descriptive representation of the body as the aspect being evaluated rather than the general individual.

A second content relevance issue that was raised in the qualitative feedback was the use of reverse-scored items. Some of the experts in this study suggested that reverse-scored items measure a different factor than positively worded items or that there is a measurement bias created by reverse-scoring leading to lower item loadings for reverse scored items than positively worded items (see Rodebaugh, Woods, & Heimberg, 2007). The use of positively worded items and reverse-scored items in the same scale has been found to be problematic and may contribute to low internal reliability among items (Hazlett-Stevens, Ullman, & Craske, 2004; Marsh, 1996a). The expert panel made specific reference to citations surrounding this issue. A concern with reverse-scored items was noted on 22 occasions by the experts. A positive comment associated with one reverse scored item "I am physically weak" (Item 18) was noted. The negatively worded items identified as confounding with other GPSCS constructs were removed (i.e., "My body is flabby" (Item 7) was a reverse-scored muscle-tone items that was removed as it was suggested to confound with the body-fat construct). Others (Items 12, 18, 22, 24, 26, 28, 30) were retained for examination in future internal structural validity studies.

Table 2.4

Summary of the Qualitative and Quantitative Data of the Sixteen Judges' Ratings of Content Relevance of the General Physical Self-Concept Scale

Construct	*Item*	Comments
Muscle-tone	1. My muscles are toned.	- maybe too general, it doesn't capture a specific facet of muscle-tone
		- (see comment under #3) I'm contrasting this with Item #1 and wondering if you are interested in whether or not muscles "appear" toned or "are" toned. I think there is a potential difference. Based on your operational definition, the items that tap in to "appearing" toned are closer to that definition than the items that say "is" or "are" toned. Just something to consider. Whether the items need to change or the operational definition needs to be more inclusive, I'm not sure. (Also, sorry if some of these comments are grammatically awkward or there are spelling mistakes. I can't actually see the type that I put in because the box is so small and the text just runs through it...)
	2. My body is firm.	- I wonder if a firm body necessarily means toned muscles. Can someone have a firm body without a lot of muscle-tone? I'd say yes.
		- The whole body? Just relevant to muscles?
		- There is no reference to tone -- I could be firm but not defined and thus Excellent or poor -- and no form of comparison
		-Not necessarily related to muscles specifically
		- perception of what firm means can vary
		- may be too vague or interpreted as 'not fat'
		- Firmness seem as function of leanness (low fat) as well as muscle-tone

(continued)

Construct	Item	Comments
	3. My muscles appear lean.	- I'm contrasting this with Item #1 and wondering if you are interested in whether or not muscles "appear" toned or "are" toned. I think there is a potential difference. Based on your operational definition, the items that tap in to "appearing" toned are closer to that definition than the items that say "is" or "are" toned. Just something to consider. Whether the items need to change or the operational definition needs to be more inclusive, I'm not sure. (also, sorry if some of these comments are grammatically awkward or there are spelling mistakes. I can't actually see the type that I put in because the box is so small and the text just runs through it...) - Appear to me (or others) - I am not sure someone would describe 'muscles' as being lean - one's body appears lean - To me lean is about lack of fat so how can muscles look lean? this could be confusing - Lean muscles appear to refer to muscles that are small in size (many would think of a ballerina). One would need to add 'lean' with tone to describe this particular type of body appearance
	4. My muscles are well defined.	- Related to muscularity rather than 'tone' - This is a little more generic - (same as #3) I'm contrasting this with Item #1 and wondering if you are interested in whether or not muscles "appear" toned or "are" toned. I think there is a potential difference. Based on your operational definition, the items that tap in to "appearing" toned are closer to that definition than the items that say "is" or "are" toned. Just something to consider. Whether the items need to change or the operational definition needs to be more inclusive, I'm not sure. (Also, sorry if some of these comments are grammatically awkward or there are spelling mistakes. I can't actually see the type that I put in because the box is so small and the text just runs through it...)

(continued)

Construct	Item	Comments
	5. My body looks muscular.	-This item does not differentiate the appearance of having lean, toned muscles from having big bulky muscles
		- This is one item that will need some specific thought because the multiple meanings of muscularity are at stake here. Just having "muscular" could mean tone or bulk depending on whose reading it and how they interpret it. I.e., this item will have some domain overrepresentation to it.
		- Yes, but not toned
		- May relate more to size than tone
		- May not be specific to 'tone', too vague maybe
		- Related to muscularity rather than 'tone'
	6. I have a large amount of muscle mass on my body.	- This items is clearly about amount or size of muscles rather than a lean, toned appearance
		- My musculature is greater than others of my gender and age?
		- Muscle mass and tone are different things and this item is focused on a "large amount of muscle mass" so...
		- This has more to do with size then definition/tone
		- It's about bulk
		- Large amount - not sure about that phrase
		- Does not necessarily imply lean or toned muscles
		- This seems more like bulk

(continued)

Construct	Item	Comments
	7. My body looks flabby.	- This item does not address muscle-tone. Also, it would be a poor negatively scored item for muscle-tone because it is possible to believe you do not appear muscular without thinking you look fat for flabby. - Be careful you are not just asking the opposite of my body is toned. What you are moving towards are bloated specifics (Catell, 19**) - I think that this and item #8 could work, but I'm not sure the flabbiness and softness is necessarily equivalent to muscularity. But it might be worth keeping to see how they hang together with the other items. In principle though I support the idea of trying to have different kinds of items (rather than basically just saying the same thing over and over, which would be my concern with items 1-5 (do you really need all 5 of them - or will participants get frustrated that you just keep asking them the same question over and over). -Reverse scored I assume. Reverse scoring tends to enter method effects into measurement that may not be entirely desirable. So, even though I've said "good match", that rating comes with a caveat. - Obviously more related to fat but you cannot look toned with lots of fat so may work - Seems to be more relevant to body-fat
	8. My body is soft.	- This item does not address muscle-tone. Also, it would be a poor negatively scored item for muscle-tone because it is possible to believe you do not appear muscular without thinking you look soft. For example, you could think you look very skinny. - Muscles not body? - Reverse scored I assume. Reverse scoring tends to enter method effects into measurement that may not be entirely desirable. So, even though I've said "good match", that rating comes with a caveat. - I think many may relate this to skin texture I would expect a gender bias to this question

(continued)

Construct	Item	Comments
Muscle-bulk	9. I have large muscles.	- Related to muscularity rather than 'tone' - This is a little more generic - (same as #3) I'm contrasting this with Item #1 and wondering if you are interested in whether or not muscles "appear" toned or "are" toned. I think there is a potential difference. Based on your operational definition, the items that tap in to "appearing" toned are closer to that definition than the items that say "is" or "are" toned. Just something to consider. Whether the items need to change or the operational definition needs to be more inclusive, I'm not sure. (Also, sorry if some of these comments are grammatically awkward or there are spelling mistakes. I can't actually see the type that I put in because the box is so small and the text just runs through it...)
	10. I have bulky muscles.	- While this item uses the same wording as the title of the construct, the matched operational definition seems problematic to me. Specifically, perceiving that one has large muscles or substantial muscle mass seems to connote a positive perception for someone who values large muscles. However, to say those muscles are, or look bulky has a more negative connotation. I believe this distinction is subtle, which is why I rated it a good match rather than fair or poor. - Bulky - might create some challenges in interpretation - (same as above) My comments would be similar to the other section - I think the operational definition needs to be more clear as to whether it is "appear" to have muscle-bulk or "has" muscle-bulk. Then the items need to reflect that choice. And that choice would influence the match I see in the items you've presented.
	11. I have huge muscles.	- (same as two above) All you are doing is asking the same question but in magnitude each time BLOATED SPECIFICS - May work to define extremes

(continued)

Constructs	Items	Comments
	12. I have small muscles.	- Good as a reverse scored item, opposite of #9 - I'm getting a bit concerned that all of the items are asking the same thing...I think you'll get a high alpha - but are the multiple items really needed doing it this way. This is a very difficult task in scale development, but I encourage you to really think about it - otherwise my fear is that you'll end up with a 20-item, 5 subscale scale with the 5 items in each subscale essentially just being repetitive. It might end with good measurement properties, but I think would add unnecessary burden to your participants - and I'm not sure you'd really be pleased with the end result of the scale (and your reviewers when you submit for publication might pick up on this and just say all you've done is repeated the same question over and over). Sorry to ramble, but based on my experience with scales I strongly recommend at least thinking about this issue as you're putting the scale together. You're putting a lot of work into this scale development, and you don't want to end up with something that someone could just as easily assess using 4 items that basically just ask a question like "I have toned muscles" for each domain.
	13. I have scrawny muscles.	- See my earlier comments on reverse scoring of items. - 'scrawny' north American term - OK as a reverse scored item. Again I think the negative evaluate tone of this wording compared to items that focus on a description of the size or mass of muscles may be problematic.
	14. I have no muscles.	- This item does not seem adequately distinct from muscle-tone. - Impossible to have no muscles. You would be just a bag of bones without them - Impossible? - See my earlier comments on reverse scoring of items. - May work as an extreme - Not possible, It isn't realistic to say "no" muscles - Everybody has "muscle'---just the extent to which it is perceived to self (and or others) as having 'size'

(continued)

Construct	Item	Comment
	15. My muscles are slender.	- Very good as a reverse scored item - Not a big fan of "slender", but this might just be a personal preference. - See my earlier comments on reverse scoring of items. As well, for some people this might be confounded with muscle-tone. - Seems more relevant to muscle-tone and also gender bias - Good match for definition, but 'slender' doesn't seem like an appropriate term - Is this like toned? - Muscles aren't often thought of as being slender
	16. I have bulging muscles.	- Sounds like strange choice of wording - bulk seems better - I think muscles could be big, but not necessarily bulging. To me, bulging includes the shape moreso than just bulk - heading it a bit towards toneness. - This does not seem entirely distinct from muscle-tone. It also may be problematic if bulging is interpreted as connoting a negative evaluation.
Muscle-strength	17. I am physically strong.	- Muscle-strength is quite different from 'tone' or 'bulk' that refer to the looks of the body. Strength refers person ability to do things (e.g., lift). It is also an objectively measurable, unlike tone. Therefore, it does not work as well as a 'self-perception' item.
	18. I am physically weak.	- Very good as a reverse scored item - See my earlier comments on problems with reverse scored items. - Opposite of 17 - Negative items are always a challenge

(continued)

Construct	Item	Comments
	19. I can lift heavy objects.	- "heavy" is open to individual interpretation - I am strong enough to lift ... - Give example of what heavy would mean - I think people might think more about injury in this case, not just pure strength. - Can be skinny but strong
	20. I cannot lift heavy objects.	- Opposite of 19. no new info will be gained here - See my earlier comments on problems with reverse scored items. - Can be limited by back pain etc - (see comment from #19) give example of what heavy would mean - (see comment from #19) "heavy" is open to individual interpretation
	21. I am physically powerful.	- May be more related to size - Power and strength are often used colloquially to mean the same thing--but some people are aware that strength is the ability to generate force while power is the ability to generate force quickly. This may result in a confound. - The knowledgeable consumer will differentiate between power and strength. Even myself, I'd say I have pretty good muscle-strength, but my self-perception of power would be much lower. - Repeat
	22. I am physically feeble.	- Reverse scored - Feeble can have different meanings - I think this would include more than just strength. - See my earlier comments on problems with reverse scored items.

(continued)

Construct	Item	Comments
	23. My muscles are strong.	- Repeat
	24. My muscles are weak.	- Reverse scored - No new info in asking this - See my earlier comments on problems with reverse scored items.
Body-fat	25. My body is fat.	- This may be interpreted more as shape than actual body-fat - Similar to strength, body-fat is objectively measurable and thus, is not dependent on self-perception only. It is different to measure body-fat (%) that to ask if individuals perception themselves fat. You might need to change this item 'fatness' as often the body-fat% has not much to with the perception of fatness
	26. My body is skinny.	- See my earlier comments on problems with reverse scored items. - Opposite of 25 - Reverse scored
	27. My body is large.	- Does not distinguish large muscles from large and fat. - I think this will get very confused with muscle-bulk. - Could be mistaken for height - Could be large without fat - Large doesn't discriminate between largeness as a consequence of adiposity and and largeness as a consequence of muscularity - Is large shape or muscularity? - Confounded with bulk - Confounded with large muscular bodies - Large is not necessarily fat - Large could be interpreted as having muscle mass (or being tall)

(continued)

Construct	Item	Comments
	28. My body is thin.	- Reverse scored - Same as skinny - See my earlier comments on problems with reverse scored items.
	29. I have a plus-sized body.	- In this case it may be OK, but body size is not always interpreted as body-fatness. - Good if this is for females only - Plus size is a north amercian term- - Not as phrase I am familiar with - Gender bias - a term not often used by males - Good match - but, men might not know how to respond to this - Perhaps gender biased - (see comment for #27)I think this will get very confused with muscle-bulk.
	30. I have a lean body.	- May be some overlap with muscle-tone - This is actually an interesting one. Maybe the leanness and body-fat scales are one and the same...(at least from a measurement perspective) - Leanness can sometimes be confounded with tone. - More relevant to body tone
	31. I am chubby.	No comments
	32. I have a plump body.	No comments

Other qualitative feedback identified terms that may not be consistently interpreted across culture and gender. Some of the questionable terms included, “slender”, “plus sized” and “toned”. The terms “slender” and “plus sized” are solely used among women according to the expert judges, therefore items using these terms should be eliminated or reworded in the instrument. Second, the term ‘toned’ is not cross-cultural. According to our expert panel, in the United Kingdom the term “henched” would be more appropriate than “toned”. Third, terms such as “feeble”, “scrawny”, “powerful”, and “having no muscles”, were considered derogatory, too extreme and/or not clearly matching the construct. The terms “feeble” and “scrawny” were considered derogatory and items using these terms were subsequently removed from the instrument. The term “powerful” is a very specific form of strength which goes beyond the specificity of the operational definition of muscle-strength, therefore this item was removed. Finally, the term ‘having no muscles’ was identified as being actually impossible, therefore this item was also removed. Collectively, the qualitative feedback from the expert panel provided informative guidance on ways to improve and/or eliminate items.

Content representativeness analysis of the general physical self-concept scale.

***Screening for discrepant raters*.** The content representativeness data were screened for discrepant raters using the methods identified in the content relevance analyses. No aberrant judges with scores considerably greater than the group of judges’ JDM scores were identified; see Table 2.5.

Content representativeness. Table 2.5 displays the judges' ratings for the muscle-tone, muscle-bulk, muscle-strength and body-fat item sets, as well as the descriptive statistics associated with each rating. The mean content representativeness rating for the item-sets were as follows; muscle-tone 3.67 (*SD*=.77), muscle-bulk 3.59 (*SD*=.65), muscle-strength 3.81 (*SD*=.75) and body-fat 3.29 (*SD*=1.21). The mean and median ratings for the four item sets ranged from "good" (3) to "very good" (4) which is an indication that the academic experts felt that the item sets adequately covered the construct of interest.

As seen in Table 2.5, the V-coefficients for muscle-tone, muscle-bulk, and strength were significant ($p < .01$), whereas the V-coefficient for body-fat was non-significant. It appears that the content representativeness of the body-fat items need to be improved whereas the content representativeness of the muscle-tone, muscle-bulk and muscle-strength item sets appear to be adequate.

Qualitative feedback of content representativeness ratings. The qualitative feedback regarding content representativeness is contained in Table 2.6. The content representativeness feedback identified areas of improvements within the item-sets. First, the experts identified specific items in the muscle-tone and body-fat item-sets that had a strong overlap. One reviewer identified that "flabby is going to crossload with adiposity". According to these comments, the reverse-scored items for muscle-tone such as "my body looks flabby" (Item 7) would be more appropriately used as a positively worded item for the body-fat construct. The likelihood of a close relationship between the reverse-scored muscle-tone items and the positively worded body-fat item resulted in the reverse-scored muscle-tone items being removed. Other feedback from the judges included making a clear distinction between the "body" and "muscle" for the muscle-tone construct. For example, the item "my body is firm" (Item 2) does not clearly refer to muscularity; whereas the item "my muscles appear firm" would improve the clarity. This particular respondent suggested that three of the items (2, 7 & 8) within the muscle-tone construct referred to the body rather than the muscle and needed to be modified to reflect muscularity.

Further feedback from the judges concerning the muscle-bulk, muscle-strength and body-fat item-sets indicated that some items within the scales were redundant and could be removed without lowering the content representativeness of the item-sets (see Table 2.6). An example of the judges' concern of redundancies among items representing the constructs was for the muscle-bulk construct. One judge suggested that "bulky, bulging, and large muscles sound similar" and may be redundant. In scale development, capturing the phenomenon of interest with a set of items that reveals the construct in different ways often includes using multiple seemingly redundant items (DeVellis, 2012). Redundancies are more prevalent in scales that tap into variables with high specificity (DeVellis, 2012). The GPSCS included constructs that have high levels of specificity such as muscle-tone, muscle-bulk, body-fat and muscle-strength therefore a certain level of redundancy would be expected.

Table 2.5

Item-Set Content Representativeness Ratings and Descriptive Statistics of the Overall Body Scale

	Expert Panel Judge Identification Number																						
Constructs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	*M*	*Mdn*	*SD*	R	Aiken'sV
Muscle-tone	5	3	3	4	3	3	4	4	3	4	4	3	3	5	3	4	5	3	3.67	3.5	.77	3	.67*
Muscle-bulk	5	3	2	4	3	3	4	4	3	3		3	4	5	4	4	4	3	3.59	4	.80	4	.65*
Muscle-strength	5	4		3	4	3	5	4	3		4	3	4	5	3	4	4	3	3.81	4	.75	3	.70*
Body-fat	5	2	2	4	4	3	3	5	2		2	4	3	5	3	4	4	1	3.29	3	1.21	5	.57
JDM	5.5	2.5	3.5†	2.5	2.5	2.5	1.5	2.5	3.5	1.5†	1.5†	3.5	.5	5.5	1.5	1.5	2.5	4.5	2.72	2.5	1.40		

Note. † Not all constructs were responded to therefore median was used as the score in the assessment of JDM. All JDM scores are within 2.7 *SD* of the mean. $^{*}p < .01$.

However it was clear that eight items were not required to cover the representativeness of these constructs. An instrument with four items per construct/subscale would be appropriate to represent muscle-tone, muscle-bulk, muscle-strength and body-fat. Items that were specifically and consistently identified as problematic were removed while other items with inconsistent concerns or no concerns were retained for further psychometric assessment. Although it was acknowledged that four items is an adequate number of items to represent each of these constructs.

Developments from the content relevance and representativeness of the GPSCS. Two researchers (the PhD candidate and her supervisor) identified the problematic items according to the quantitative data and qualitative feedback. The qualitative feedback provided by the judges was used to make modifications to problematic items and also to select those items that needed to be eliminated. The modifications to the items were first developed independently by the two researchers. The final version of the modified items to be included in the second study of this dissertation was achieved through discussion to consensus between the two researchers. Table 2.7 lists the original items and the final list of items (including modifications) that comprised the item set to be used in the subsequent validity study.

The comments from the judges included feedback that would improve the clarity of the operational definitions. This feedback was used to modify the operational definitions according to qualitative feedback received from the judges. Muscle-tone, muscle-bulk, and body-fat operational definitions were modified to consistently refer to the ‘degree’ of muscle or body-fat, and secondly refer to the appearance of the body/muscle rather than how the body/muscle ‘is’. Muscle-tone, muscle-bulk and muscle-strength items were specified to refer to the muscle rather than the body as a whole. A list of the original operational definitions and updated operational definitions are included in Table 2.8.

Discussion

The purpose of this study was to examine the content relevance and content representativeness of items intended to measure muscle-tone, muscle-bulk, muscle-strength and body-fat. The feedback and ratings provided by the academic expert judges were used to make modifications to the items prior to further content relevance and representativeness assessments. Overall, the item content relevance could be improved by considering the feedback from the judges and incorporating the expert panel’s suggestions. Good item-set content representativeness was found according to the analyses.

Muscle-tone

The academic experts identified some concern associated with the negatively worded (i.e., reverse scored) items in this study. Some measurement experts in the social sciences suggest that scales should include an equal number of positively and negatively worded items representing a single construct (DeVellis, 1991; Pedhazur & Schmelkin, 1991). The purpose of incorporating both positively and negatively worded items to represent a single construct is to prevent acquiescence, which is the tendency for participants to agree with a statement when they may be unsure or ambivalent (Nunnally, 1978).

Table 2.6

Qualitative Assessment of Content Representativeness of the Overall Body Scale

	Expert Responses			
	Additional items		Redundant items	
Construct	Yes	Comments	Yes	Comments
Muscle-tone	6	-There could be cultural specific terms to be used in certain context or population. UK kids refer to being toned as 'hench'! But of course this is difficult to incorporate. Just wanted to raise this as a general issue. - Perhaps related to body being sculpted - no time to think of any at the moment - I'm not sure why you use "firm" in the construct characterization but not in any item content - in many cases you have used a thesaurus like approach, in many cases simply replacing a single word with a Synonym or an Antonym. I am not sure this is the best way to capture the universe of the construct - you could talk about being 'ripped' or 'cut' but that might not have meaning to some people - I wonder if you need to be clear when you say "looks" muscular (for example) who the audience is. Who I say is "looking" could make a big difference. For example, if I interpret it as "others" looking, I might have a very different response than if I contextualize it as what I see when I look at myself. This is relevant to the scale as a whole. Who is the audience? If it is self-perception, then maybe it needs to be clarified that it is the self who is looking - and not others (unless you want to detect what they would report others would see in oneself).	12	- 2 expert suggested removing item 2, 5 experts suggested removing Item 7, and 6 experts suggested removing item 8. - Items 1 and 2 appear the best here - 5 out of 8 items are about the body appearance not muscle appearance - I would like to see "body" versus "muscles" better differentiated or defined operationally. For example, at least questions 2, 7, 8 can be changed to reflect muscles rather than body. - The first item is the same as the construct. - I have reservations about using reverse scored items. As well, "flabby" is going to cross load with adiposity items. - I think I covered this in a earlier section, but 2 suggestions: (a) given the item overlaps, I strongly encourage you have no more than 3-4 items per scale in the final result. With the similarities among items, there seems no need to have an 8 item scale. (b) the final 2 items overlap alot with the body-fat scale (so a choice will need to be made where they really fit - or if you just have one scale between those 2 dimensions).

(continued)

		Expert Responses		
		Additional items		Redundant items
Construct	Yes	Construct	Yes	Construct
Muscle-bulk	2	- the items may be worded such that they reflect terms more often used with males than females - I wonder if you need to be clear when you say "looks" muscular (for example) who the audience is. Who I say is "looking" could make a big difference. For example, if I interpret it as "others" looking, I might have a very different response than if I contextualize it as what I see when I look at myself. This is relevant to the scale as a whole. Who is the audience? If it is self-perception, then maybe it needs to be clarified that it is the self who is looking - and not others (unless you want to detect what they would report others would see in oneself). - In many cases you have used a thesaurus like approach, in many cases simply replacing a single word with a Synonym or an Antonym. I am not sure this is the best way to capture the universe of the construct - Should #12 be reverse coded?	14	- experts suggested that the following item could be removed item 10 (N=1), item 13 (N=4), item 16 (N=2), item 15 (N=1), item 14 (N=4), item 12 (N=1) and item 11 (N=1). - items 10,11,12, 13 and 16 are redundant - Items 9 and 10 appear the best here - the terms 'bulky' and 'scrawny' are going to be tough to use both because of possible interpretational differences for males and females, some people may not know the term scrawny, and "bulk" may infer a negative connotation (i.e., my muscles are bulky and therefore prevent me from doing tasks...) - I also have reservations about reverse scored items here. - bulky, bulging, and large muscles sound very similar - See comments in last section, which I think are generally relevant here as well (except for (b))
Muscle-strength	6	- avoid negatively worded items (see Fletcher and Hattie) - Does fatigue fit in her? - In all these there is too much "bloated specifics" - items are quite extreme, almost like polar opposites? is this the intent? Semantic differential?	11	- one expert suggested that 18, 20,22,23 and 24 are redundant - experts suggested that item 18 (N=1), item 19 (N=1), item 20 (N=2), item 23 (N=1) and item 24 (N=1) may be removed - The reverse scored items are opposites of the other items, so seem redundant.

(continued)

		Expert Responses		
		Additional items		Redundant items
Construct	Yes	Construct	Yes	Construct
		- as strength refers to the muscles ability to contract, you might need more things that refer to the body's ability to things maximally, few times. At moment, your reference to strength is also quite gendered as men usually attribute strength to upper body. Can women (or others) conceive themselves strong even if they can lift things?		- You appear to have four sets of opposites, so one could interpret that you have four redundant items - The power item may offer a confound to the measurement. Again I have reservations about reverse scored items. - I think it is questionable to use exact opposites in the same scale - remove negatively worded item Not redundant, but items 21 and 22 may be getting at a different aspect (power is not just strength)
Body-fat	7	- It not so much that you would need to add items, but you do really measure body-fat here, but people's perceptions of 'fatness' - Again, is the intent to use semantic differential format? - Could include an item or two about having a lot of body-fat to pinch, or being able to see a lot of fat on one's body - Items about having a high percent of body-fat; visible fat - These should be cut by half -- getting rid of the bloatedness - "My body is large" could be fat or it could be muscularity and hence	15	- Experts identified that the following items may be eliminated item 26 (N=3), item 27 (N=5), item 28 (N=3), item 29 (N=3), item 30 (N=1) and item 31 (N=2). - many of the items developed for this subscale are opposites of each other (fat/skinny; large/thin) - plus-size is a bit rubbish - item 29 might be difficult to assess due to possible gender differences in interpretation and the fact that some people might not know what a plus size body is. - This scale is the one most likely to reflect gender bias. The large body could be a a=large muscular body, plus size is an expression most common to female - Items related to shape more than fat - The reverse scored items may create problems. AS well, the adiposity and muscle-tone are confounded somewhat in the item set in my opinion.

(continued)

		Expert Responses		
		Additional items		Redundant items
Construct	Yes	Construct	Yes	Construct
		- I wonder if you need to be clear when you say "looks" muscular (for example) who the audience is. Who I say is "looking" could make a big difference. For example, if I interpret it as "others" looking, I might have a very different response than if I contextualize it as what I see when I look at myself. This is relevant to the scale as a whole. Who is the audience? If it is self-perception, then maybe it needs to be clarified that it is the self who is looking - and not others (unless you want to detect what they would report others would see in oneself).		- #27 and #29 especially because I think someone high in muscle-bulk would rate themselves high on those items, even if they have little fat.

Other researchers suggest that including positively and negatively worded items introduces a method effect which may result in two factors being measured while assessing a single construct (Marsh, 1996b). Hazlett-Stevens, Ullman and Craske (2004) tested the Penn State Worry Questionnaire's single construct of worry and found a method effect that resulted in two separate factors; one factor that reflected the degree in which worry is present and one factor that reflected the degree to which worry is absent. Marsh (1996a) found this similar method effect when testing Rosenberg's (1965) Global Self Esteem Scale. Recently, DeVellis (2012, p.84) reported that any benefits associated with using positively and negatively worded items within a scale is out-weighed by the consistent poor performance of the negatively worded items. This may be particularly important for constructs such as self-perceptions of muscle-tone, muscle-bulk, and body-fat that do not have a distinct identifiable opposite. The combination of literature and comments from the expert panel supported the removal of the reverse scored items in the muscle-tone subscale for the updated version of the GPSCS.

A second issue raised by the expert panel was that Item 5 (My body looks muscular.) was considered too general and did not distinguish between perceptions of muscle-tone and muscle-bulk. The proposed use of this instrument was to distinguish between muscle-tone and muscle-bulk, therefore it was deemed necessary to remove this item due to a lack of clarity.

The third concern of many expert-panel members was the lack of precision of the item wording for some of the muscle-tone items according to the operational definition. More specifically, the operational definition for muscle-tone stated, “the self-perception of the degree to which muscles look lean, firm and defined” while some items used the terms *appearing* muscular and others referred to as *being* muscular. Another point of item clarity for the muscle-tone and muscle-bulk item sets was the distinction between the assessment being of the *body* or the *muscle*. For example “my body is firm” (Item 2) could be more explicitly stated as “my muscles are firm” to improve the clarity of the meaning of the item. In order to take into consideration conceptual clarity of the muscle-tone construct the item could be improved by rephrasing to say “my muscles are firm”. The traditional physical self-concept definition is a “person’s perceptions, thoughts and feelings about his or her body” (Schilder, 1950, p.11) and is modeled according to Shavelson’s (1976) model which refers to a descriptive and evaluative representations of the individual in a particular dimension of the self.

Therefore, in order to maintain the conceptual clarity of the operational definition, as well as consistency with the traditional physical self-concept definition, the items were all modified to address the appearance of the muscles rather than an assessment of their actual state.

The content representativeness of the muscle-tone items was judged to be 'very good' as reflected by a significant validity coefficient. Although the content representativeness results are supportive of the items, expert comments suggested that fewer items could be used to cover the breadth of the muscle-tone construct. Shavelson's (1976) conception of self-concept includes a hierarchy of general self-concept at the apex, followed by academic, social, emotional and physical self-concept on the 2nd tier. Self-perceptions of muscle-tone, muscle-bulk, muscle-strength, and body-fat would be included on the 3rd tier of the hierarchy under the construct of physical self-concept. The constructs of self-perception of muscle-tone, muscle-bulk, muscle-strength and body-fat are very specific constructs that do not include great breadth, having eight items representing these construct result in potentially redundant items (see Clark & Watson, 1995). The original pool of 8-items per construct was considerably more items than was necessary for the final item-set due to the narrow level of specificity of the constructs. Optimally the constructs will include four items per construct in order to cover the breadth of the construct without being redundant (Clark & Watson, 1995). The final version of the GPSCS was planned to include four items per construct/subscale in order to minimize participant burden while maintaining the psychometric soundness of the instrument.

Table 2.7

Original GPSCS Items and Modified GPSCS Items for Study 2

	GPSCS Items	
Construct	Original Items	Items for Study 2
Muscle-tone	1. My muscles are toned.	1. My muscles appear toned.
	2. My body is firm.	2. My muscles appear firm.
	3. My muscles appear lean.	3. My muscles appear lean.
	4. My muscles are well defined.	4. My muscles appear well defined.
	5. My body looks muscular.	(removed)
	6. I have a large amount of muscle mass on my body.	(removed)
	7. My body looks flabby.	(removed)
	8. My body is soft.	(removed)
Muscle-bulk	9. I have large muscles.	5. I appear to have large muscles.
	10. I have bulky muscles.	6. My muscles appear bulky.
	11. I have huge muscles.	7. My muscles appear to be huge
	12. I have small muscles.	8. My muscles appear small.
	13. I have scrawny muscles.	(removed)
	14. I have no muscles.	(removed)
	15. My muscles are slender.	(removed)
	16. I have bulging muscles.	9. I appear to have bulging muscles.

(continued)

Construct	GPSCS Items	
	Original Items	Items for Study 2
Muscle-strength	17. I am physically strong.	10. I am physically strong.
	18. I am physically weak.	11. I am physically weak.
	19. I <u>can</u> lift heavy objects.	12. I am <u>strong enough to</u> lift heavy objects.
	20. I cannot lift heavy objects.	(removed)
	21. I am physically powerful.	(removed)
	22. I am physically feeble.	13. I am physically feeble.
	23. My muscles are strong.	14. My muscles are strong.
	24. My muscles are weak.	15. My muscles are weak.
Body-fat	25. My body <u>is</u> fat.	16. My body <u>appears</u> fat.
	26. My body <u>is</u> skinny.	17. My body <u>appears</u> skinny.
	27. My body is large.	(removed)
	28. My body <u>is</u> thin.	18. My body <u>appears</u> thin.
	29. I have a plus-sized body.	(removed)
	30. I <u>have a</u> lean body.	19. My body <u>appears</u> lean.
	31. I <u>am</u> chubby.	20. I <u>appear</u> chubby.
	32. I <u>have a</u> plump body.	21. I <u>appear to</u> have a plump body.

Note. Words are underlined to highlight modifications made to the items.

Table 2.8

Original GPSCS Operational Definitions and Modified GPSCS Operational Definitions for Study 2

	GPSCS Operational Definitions	
Constructs	Original	For Study Two
Muscle-tone	The self-perception of the degree to which muscles look lean, firm and defined.	The self-perception of the degree to which muscles <u>appear</u> lean, firm and defined.
Muscle-bulk	The self-perception of the <u>size and mass of muscles</u>.	The self-perception of the <u>degree to which muscles appear large and bulky in size</u>.
Muscle-strength	The self-perception of the presence or absence of strength and the ability to lift heavy objects.	The self-perception of the presence or absence of <u>muscle-strength</u> and the ability to lift heavy objects.
Body-fat	A self-perception <u>relating to the amount</u> of body fat.	The self-perception <u>of the degree of fat that appears on the body.</u>

Note. Words are underlined to highlight modifications made to the operational definitions.

Muscle-bulk

The content relevance and representativeness of muscle-bulk was considered very good by the expert panel. The main feedback from the judges was the same as for muscle-tone: concerns about the reverse-scored items; and concern about the lack of item specificity with reference to the ‘appearance’ of the muscle-bulk. These two concerns were addressed similarly to the muscle-tone items and resulted in the elimination of three items and the rewording of 5 items. One reverse-scored item was retained for further testing with the next group of study participants to confirm these suggestions with a new sample of experts.

Muscle-strength

Overall, the expert panel agreed the muscle-strength item-set had good representativeness of the construct. This construct was found to have better content validity than the other constructs being tested which resulted in fewer modifications being made to these items. Perceptions of muscle-strength have previously been assessed in PSC measures (i.e. PSDQ, Marsh, 1996; PSPP, Fox & Corbin, 1989) and are a more concrete and explicable construct than appearance constructs such as muscle-tone and muscle-bulk. Objective assessments of strength are common among the general population in gym classes, fitness tests or through self-assessments which provides one with specific feedback about their strength.

Although the muscle-strength items generally demonstrated good content relevance and representativeness, it was possible to improve the item-set according to feedback from the expert panel. For example, “I am physically powerful” (Item 21), included a very specific form of strength that does not take into consideration the breadth of the operational definition of muscular strength. If the goal of the strength construct was to explore specific forms of the physical self-concept of strength this would be a good item. This construct was intended to develop a general measure of self-perceptions of physical self-concept to contrast with the self-perception of muscle-tone and muscle-bulk therefore this item was too specific. Due to the over specificity of the item relative to the operational definition of muscle-strength, this item was removed.

As mentioned previously with the muscle-tone and muscle-bulk constructs, the experts identified some concern with the reverse-scored items in the muscular strength item-set. The use of reverse scored items with the muscle-strength construct is different than the appearance oriented constructs, as the operational definition explicitly identifies the absence of strength, and muscle-strength is a simpler concept to understand than muscle-tone and muscle-bulk. Muscle-strength is a simpler construct for most people to understand and respond to as it can be easily measured while lifting weights or participating in activities of daily living, such as carrying groceries. A continuum of strength ranging from *weak* to *strong* is a well accepted and empirically supported (Marsh, 1996), whereas constructs such as muscle-tone have not been well distinguished and are more complex. One reverse-scored item was removed while three remained in the item pool to undergo further testing with the fitness experts and in factor analysis studies to assess for any ‘method’ effect. A ‘method’ effect in this situation is when items representing two poles of a hypothesized construct actually measure two distinct constructs rather than a difference in the degree of the presence of the construct (Clark & Watson, 1995).

Item 20, ‘I can lift heavy objects’ was also a concern with the experts. The interpretation of what ‘heavy’ would include was identified as an issue among experts. Some suggested including an absolute measure, such as ‘I can lift a 20-kilogram weight’. This item is problematic for two reasons. Given that individual factors such as gender, body size, and type of athletic experience result in different conceptions of what a heavy weight weights, it might not be possible to agree on a specific weight to include in the item. Secondly, lifting a weight once is a specific form of strength rather than a general measure of strength. One may consider them self a strong person because they can do 10 chin ups, hold a plank position for 2 minutes, or bench press their body weight. This item was marked as being problematic according to the academic experts but retained for further testing among fitness experts as this is their area of expertise.

Body-fat

The items representing the fourth construct, body-fat, generally demonstrated good median and mean content relevance scores, but the group of experts as a whole did not endorse content relevance scores high enough to claim content-relevance according to the non-significant V-coefficients. Although the assessment of self-perception of body-fatness has been around for decades, it remains a complex construct. Objective assessments of body-fat are rarely conducted. Body composition assessment techniques such as skin-fold measurements or dual-energy x-ray absorptiometry (DEXA) scanning can provide a percent body-fat and a categorical identification in relation to health. The different reference points of *a good amount of body-fat* varies from person to person. Some people desire a “healthy” amount of body-fat; some aspire to have bodies like the male and female models in fitness magazines; some want to look like runway models; and others want to have less body-fat than those in their social network (Groesz, Livine & Murnen, 2002; Olgle & Damhorst, 2005). The different conceptualization of the evaluative component of how much body-fat is acceptable lacks is not clear.

More specifically, Item 27, ‘I have a large body’, was identified as being too general and open to multiple interpretations. This item could be interpreted as having a lot of body-fat, being tall or having lots of muscle mass. Therefore due to a lack of specificity and clarity, this item was removed. Next, Item 29, ‘I have a plus-sized body’, was identified as being problematic for two reasons. First, plus-sized is a North American term therefore this item would have low ecological validity elsewhere. Second, plus-sized is a term used to describe females not males, which would further limit the scope of this instrument. For these reasons, this item was also removed.

Self-perceptions of the appropriate amount of fat may vary according to: sex, age, type of athletic experience, social networks, personality, etc. Although there are currently instruments available to assess perceptions of body-fat, many lack rigorous psychometric testing and often examine weight and body-fatness under the same construct of self-perception or body dissatisfaction which is less specific than the GPSCS and cannot necessarily distinguish between self-perceptions of body-fat and muscle-bulk.

Summary

Theoretically, self-concept organizes self-knowledge and guides how the self is perceived by the individual, focuses motivation, regulates emotion and guides social interaction (Epstein, 1973). More specifically, it is possible that perceptions of the physical self can be used to guide exercise behaviour. This study introduced two new constructs of physical self-concept, muscle-bulk and muscle-tone, that may relate to exercise participation. The constructs of body-fat and muscle-strength were included in the GPSCS to assure that the muscle-tone and muscle-bulk constructs/subscales are distinguishable from them. The body-fat construct was included in the GPSCS as body-fat may be closely related to muscle-tone. In order to see muscle definition and tone, one cannot have a lot of body-fat covering the muscle therefore body-fat was included in this scale. Secondly, muscle-strength was included to distinguish from the functional aspect of muscularity. The presence of bulky muscles and muscle-strength were expected to be related but distinct as well, as muscle-strength may still be observed with the absence of bulky muscles.

The expert panel supported these four constructs and items representing them. Establishing validity evidence is an ongoing process of accumulating evidence to inform the interpretation of test scores and their use (Messick, 1995). This was an initial study which assessed the content relevance and representativeness of the proposed GPSCS items with a panel of academic experts.

The inclusion of physical self-concept experts from around the world as participants on the expert panel was critical in the identification of item limitations across cultures. Some terms within the GPSCS items were identified as being foreign and unfamiliar in certain cultures which resulted in poorer content relevance and content representativeness scores for some items (i.e., the North American term ‘toned’ is not used in the United Kingdom where the underlying idea is referred to as henched). The cultural specificity of some terms used in the GPSCS identified that this instrument may not appropriate for use in the United Kingdom. Consequently, the scope of the target population for this instrument was limited to North Americans. Second, the expert panel contained academics who may or may not have experience talking with young adults about muscle-tone, muscle-bulk, body-fat, and muscle-strength. As such, the expert judges may be unaware of the terms use by young adults to describe these concepts. This instrument has been constructed to assess self-perceptions of muscle-tone, muscle-bulk, body-fat and muscle-strength among young adults. There is a need, therefore, to consult content experts from the practical domain as well as theoretical experts. These ‘practical’ experts will provide a unique and critical role in developing items for assessment that is relevant to the population of interest which is where the ideas came from in the first place.

In summary, the purpose of this study was to examine the content relevance and content representativeness of items proposed to assess muscle-tone, muscle-bulk, muscle-strength and body-fat according to an academic expert review process. Overall, good content representativeness was observed while the content relevance of the GPSCS may be improved. The feedback and results from this study were used to refine the item set for further psychometric evaluation.

References

Aiken, L.R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*, 131-142. doi:10 .1177/0013164485451012

Bless, C., & Higson-Smith, C. (2000). *Fundamentals of social research methods: An African perspective* (3rd ed.). Cape Town, ZA: Juta and Company.

Blood, S.K. (2005). *Body work: The social construction of women's body image.* London, GB: Routledge.

Butler, J. C., & Ryckman, R. M. (1993). Perceived and ideal physiques in male and female university students. *Journal of Social Psychology, 133*, 751-752. doi:10.1080/00224545.1993.9713935

Cafri, G., & Thompson, J. K. (2004). Evaluating the Convergence of Muscle Appearance Attitude Measures. *Assessment, 11*, 224-229. doi:10.1177 /1073191104267652

Cafri, G., & Thompson, J. K. (2007). Measurement of the muscular ideal. In J.K. Thompson & G. Cafri (Eds.), *The muscular ideal: Psychological, social, and medical perspectives* (pp. 107-120). doi:10.1037/11581-005

Cafri, G., van den Berg, P., & Thompson, J. K. (2006). Pursuit of muscularity in adolescent boys: Relations among biopsychological variables and clinical outcomes. *Journal of Clinical Child and Adolescent Psychology, 35*, 283-291. doi:10.1207/s15374424jccp3502_12

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*, 309-319. doi:10.1037/1040 -3590.7.3.309

Crocker, L. (1997). Assessing content representativeness of performance assessment exercises. *Applied Measurement in Education*, *10*, 83-95. doi:10.1207/s153248ame1001_5

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Belmont, CA: Wadsworth.

DeVellis, R. F. (1991). *Scale development: Theory and application*. London, GB: Sage.

DeVellis, R. F. (2012). *Scale development: Theory and application (3rd ed.)*. London, GB: Sage.

Dunn, J. G. H., Bouffard, M., & Rogers, W. T. (1999). Assessing item content-relevance in sport psychology scale construction research: Issues and recommendations. *Measurement in Physical Education and Exercise Science*, *3*, 15-36. doi:10.1207/s15327841mpee0301_2

Dworkin, S. (2003). A women's place is in the …cardiovascular room? Gender relations, the body, and the gym. In A. Bolin & J. Granskoog (Eds.), *Athletic intruders: Ethnogrphic research on women, culture and exercise* (pp.131-158). Albany, NY: State University of New York Press.

Epstein, S. (1973). The self-concept revisited: Or a theory of a theory. *American Psychologist*, *28*, 404-416. doi:10.1037/h0034679

Evans, R. R., Cotter, E. M., & Roy, J. L. P. (2005). Preferred body type of fitness instructors among university students in exercise classes. *Perceptual and Motor Skills, 101*, 257-266. doi:10.2466/PMS.101.5.257-266

Fitzpatrick, A. R. (1983). The meaning of content validity. *Applied Psychological Measurement, 7*, 3-13. doi:10.1177/014662168300700102

Fox, K. R. (1990). *The Physical Self-Perception Profile manual.* Dekalb, IL: Northern Illinois University Office of Health Promotion.

Fox, K. R., & Corbin, C. B. (1989). The Physical Self-Perception Profile: Development and preliminary validation. *Journal of Sport & Exercise Psychology, 11*, 408-430. Retrieved from http://journals.humankinetics.com/jsep

Frew, M., & McGillivray, D. (2005). Health clubs and body politics: Aesthetics and the quest for physical capital. *Leisure Studies, 24*, 161-175. doi:10.1080/0261436042000300432

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology (3rd ed.)*. Toronto, ON: Allyn and Brown.

Gotwals, J. (2006). *Developing a perfectionism measure in sport* (Unpublished doctoral dissertation). University of Alberta, Edmonton, AB.

Gotwals, J. K., & Dunn, J. G. H. (2009). A multi-method multi-analytic approach to establishing internal construct validity evidence: The Sport Multidimensional Perfectionism Scale 2. *Measurement in Physical Education and Exercise Science, 13*, 1-22. doi:10.1080/10913670902812663

Groesz, L. M., Levine, M. P., & Murnen, S. K. (2002). The effect of experimental presentation of thin media images on body satisfaction: A meta-analytic review. *International Journal of Eating Disorders, 31,* 1-16. doi:10.1002/ eat.10005

Gruber, A. J. (2007). A more muscular female body ideal. In J.K. Thompson & G. Cafri (Eds.), *The muscular ideal: Psychological, social, and medical perspectives* (pp. 217-234). doi:10.1037/11581-011

Hambleton, R. K. (1978). On the use of cut-off scores with criterion-referenced tests in instructional settings. *Journal of Educational Measurement, 15,* 277-289. doi:10.1111/j.1745-3984.1978.tb0075.x

Hambleton, R. K. (1980). Contributions to criterion-referenced testing technology: An introduction. *Applied Psychological Measurement, 4,* 421-424. doi:10.1177/014662168000400401

Hambleton, R. K. (1984). Validating the test scores. In R.A. Berk (Ed.), *A guide to criterion-referenced test construction* (pp. 199-230). Baltimore, MD: John Hopkins University Press.

Haravon, C. L. (2002). Working out contradictions: Feminism and aerobics. *Journal of Sport and Social Issues, 26,* 85-109. doi:10.1177/019372350 2261006

Haynes, S. N., Richards, D. C., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*, 283-247. doi:10.1037//1040 - 3590.7.3.238

Hazlett-Stevens, H., Ullman, J. B., & Craske, M. G. (2004). Factor structure of the Penn State Worry Questionnaire: Examination of a method factor. *Assessment*, *11*, 361-370. doi:10.1177/1073191104269872

Hellsten, L. M. (2005). *The development and validation of a time management scale for exercise adoption, participation, and adherence* (Unpublished doctoral dissertation). University of Alberta, Edmonton, AB.

Jacobi, L., & Cash, T. F. (1994). In pursuit of the perfect appearance: Discrepancies among self-ideal perceptions of multiple physical attributes. *Journal of Applied Social Psychology*, *24*, 379-396. doi:10.1111/j.1559-1816.1994.tb00588.x

Kyrejto, J. W., Mosewich, A. D., Kowalski, K. C., Mack, D. E., & Crocker, P. R. E. (2008). Men's and women's drive for muscularity: Gender differences and cognitive and behavioral correlates. *International Journal of Sport and Exercise Psychology*, *6*, 69-84. doi:10.1080/1612197X.2008.9671855

Lenart, E. B., Goldberg, J. P., Bailey, S. M., Dallal, G. E., & Koff, E. (1995). Current and ideal physique choices in exercising and non-exercising college women from a pilot athletic image scale. *Perceptual and Motor Skills*, *81*, 831-848. doi:10.2466/pms.1995.81.3.831

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, *35*, 382-385. doi:10.1097/00006199-198611000-00017

Markula, P. (1995). Firm but shapely, fit but sexy, strong but thin: The postmodern aerobicized female body. *Sociology of Sport Journal*, *12*, 424-453. Retrieved from http://journals.humankinetics.com/ssj

Marsh, H. W. (1996a). Physical Self-Description Questionnaire: Stability and discriminant validity. *Research Quarterly for Exercise and Sport, 67*, 249-264.

Marsh, H. W. (1996b). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70*, 810-819. doi:10.1037/0022-3514.70.4.810

Marsh, H. W ., Richards, G. E., Johnson, S., Roche, L., & Tremayne, P. (1994). Physical self-description questionnaire: Psychometric properties and a multitrait-multimethod analysis of relations to existing instruments. *Journal of Sport & Exercise Psychology, 16*, 270-305. Retrieved from http://journals.humankinetics.com/jsep

Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York, NY: MacMillan.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performance as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749. doi:10.1037// 0003-066X.50.9.741

Mosewich, A. D., Vangool, A. B., Kowalski, K. C., & McHugh, T. F. (2009). Exploring women track and field athletes' meanings of muscularity. *Journal of Applied Sport Psychology, 21*, 99-115. doi:10.1080/1041 3200802575742

Nunnally, J. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Ogle, J. P., & Damhorst, M. L. (2005). Critical reflections on the body and related sociocultural discourses at the midlife transition: An interpretive study of women's experiences. *Journal of Adult Development, 12,* 1-18. doi:10.10 07/s10804-005-1277-2

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.

Ridgeway, R. T., & Tylka, T. L. (2005). College men's perceptions of ideal body composition and shape. *Psychology of Men and Masculinity, 6*, 209-220. doi:10.1037/1524-9220.6.3.209

Rodebaugh, T. L., Woods, C. M., & Heimberg, R. G. (2007). The reverse of social anxiety is not always the opposite: The reverse-scored items of the Social Interaction Anxiety Scale do not belong. *Behavior Therapy*, *38,* 192-206. doi:10.1016/j.beth.2006.08.001

Rogers, W. T. (2010). *Notes for a lecture on construction and analysis of survey questionnaires and attitude scales.* University of Alberta, Edmonton, AB.

Rosenberg, M. (1965). *Society and the adolescent child.* Princeton, N J: Princeton University Press.

Schilder, P. (1950). *The image and appearance of the human body*. New York, NY: International University Press.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*, 407-441. doi:10.2307/1170010

Thompson, J. K., & Cafri, G. (Eds.). (2007). *The muscular ideal: Psychological, social, and medical perspectives*. doi:10.1037/11581-000

Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for establishing content validity. *Western Journal of Nursing Research, 25,* 508-518. doi:10.1177/0193945903252998

Study 2

Assessing the Content Validity of the Perception of Muscle-tone, Muscle-bulk, Muscle-strength and Body-fat Items Among an Expert Panel of Fitness Professionals

The first study of this dissertation examined the content relevance and representativeness of the proposed muscle-tone, muscle-bulk, muscle-strength and body-fat items according to the operational definitions among academic experts. Modifications to the original items were made according to the feedback from the experts. Additionally, the wording of the operational definitions was slightly adjusted taking into consideration the judges' comments on the items to ensure clarity and consistency of language use. For example, operational definitions and all items for muscle-tone, muscle-bulk and body-fat were edited to consistently indicate that the construct represents the "appearance", not necessarily the actual physical characteristic. The focus of the present paper is to establish content validity evidence for the current items, and identify areas that need modification to improve the GPSCS. The expert panel from the first study consisted of academics who had published peer-reviewed research in the area of PSC. The academic experts' knowledge of the PSC literature and psychometrics provided insight into the relevance and representativeness of the items according to the operational definitions. Although the opinions of academic experts play an imperative role in instrument construction and development, other types of experts are able to provide further assessments on the relevance and representativeness of the items (see Dunn, Bouffard, & Rogers, 1999).

As a result of the cultural sensitivity identified in Study 1 surrounding muscle-tone and muscle-bulk, the instrument being developed is intended to assess perceptions of muscle-tone, muscle-bulk, body-fat and muscle-strength among young, North American adults. Physical appearance is a salient topic among North American adults. Indeed there is a whole industry including magazines (e.g., Shape, Muscle and Fitness) that specifically cater to these concerns, and it is a dominant topic in other forms of popular media (e.g., Vogue, Chatelaine), in clothing (e. g., Lulu Lemon), not to mention garments intended to improve the appearance of the body (e. g., Spanx). One context where people specifically seek to both discuss and address their physique-related concerns is with fitness professionals and personal trainers. The knowledge acquired by fitness experts, such as personal trainers, as a result of conversations with clients make them PSC experts through practical experience. Therefore, the purpose of this study was to use fitness professionals in the expert-review process to examine the content relevance and content representativeness of the proposed muscle-tone, muscle-bulk, muscle-strength and body-fat items.

Method

Participant Characteristics

A convenience sample of nationally certified fitness experts was recruited. Participants included fitness experts possessing a post secondary education, professional fitness certification through the Canadian Society for Exercise Physiology, and experience working with clients in either one-on-one or group settings. The expert panel included 17 fitness professionals (6 males and 11 females).[2] Judges held an undergraduate degree (*n* = 17) or a Masters degree (*n* = 11) and were certified as fitness leaders (*n* = 2), personal trainers (*n* = 3), exercise physiologists (*n* = 10) and/or strength and conditioning coaches (*n* = 2).

Measures

Demographics. The first part of the survey gathered demographic information which included sex, highest degree of education, current fitness credentials, and a description of current health and fitness work. See Appendix E, Part 1 for a copy of the demographic questions.

Content relevance of the general physical self-concept scale (GPSCS). The second section of the survey assessed the content relevance of the general muscle-tone, muscle-bulk, muscle-strength and body-fat items according to the proposed matching construct (Gotwals, 2006; Hellsten, 2005). The same procedures employed in Study 1 were used to judge the degree of match between the items and the constructs they were intended to measure, and to record any comments or concerns pertaining to each item. The fitness experts were presented with four muscle-tone items, five muscle-bulk items, six muscle-strength items and six body-fat items. A total of 21 items were judged within this section (See Table 3.1 for the GPSCS items).

---

[2]Seventeen out of 19 fitness experts who indicated an interest in the study completed the survey. One expert was unable to complete the survey package due to a medical emergency and a second expert was unable to schedule a session prior to the data collection deadline.

Table 3.1

GPSCS Items According to Construct

Construct	GPSCS Items for Study 2
Muscle-tone	1. My muscles appear toned.
	2. My muscles appear firm.
	3. My muscles appear lean.
	4. My muscles appear well defined.
Muscle-bulk	5. I appear to have large muscles.
	6. My muscles appear bulky.
	7. My muscles appear to be huge
	8. My muscles appear small.
	9. I appear to have bulging muscles.
Muscle-strength	10. I am physically strong.
	11. I am physically weak.
	12. I am strong enough to lift heavy objects.
	13. I am physically feeble.
	14. My muscles are strong.
	15. My muscles are weak.

(continued)

Construct	GPSCS Items for Study 2
Body-fat	16. My body appears fat.
	17. My body appears skinny.
	18. My body appears thin.
	19. I body appears lean.
	20. I appear chubby.
	21. I appear to have a plump body.

Content representativeness of the GPSCS. The third section of the survey included content representativeness questions. The fitness experts were asked to evaluate the degree to which the group of items as a whole represented all identifiable aspects of the target construct (Dunn et al., 1999). The survey package was organized as previously described in Study 1 with one additional section. The additional section included a place where words or terms fitness experts or clients used to communicate the constructs of interest could be listed to be taken into consideration for additional items or modifications to current items. See Appendix E, Part 3 for the complete GPSCS content representativeness survey that were given to the panel of fitness experts.

Procedures

Ethical approval for the study was provided by an institutional research ethics board. The judges were sent an email that contained a very brief description of the study, identified why they were selected for this study, informed them that the survey should take approximately 30 minutes, and invited them to participate in the study by replying to the email and informed them that the survey needed to be completed within three weeks from the time the original email was sent. If the participants were interested they were to email the researcher a time when they would be available to complete the survey package in person with the researcher. See Appendix F for the study invitation sent via email and Appendix G for the study information letter.

The fitness experts completed a hard copy of the survey package in the presence of the researcher. The survey was presented section by section in order to provide the fitness experts with written and oral instructions at the start of each section. The researcher stayed in the room for the completion of the survey package to provide participants with an opportunity to ask questions at any time. Each participant's consent was implied by the overt action of completing the survey. Upon completion of the survey, the participants were asked if they had any questions about the survey or the study, and thanked for their participation.

Data were then entered into SPSS 18 and analyzed. Descriptive analysis was conducted to examine the demographics of the fitness experts. The data set was screened for missing data. Participants who responded to a minimum of 90% of the items within each section of the survey package were included in the data analysis of that section. The methods used to screen for discrepant raters, conduct the item fit analysis and calculate the item-content validity coefficient were the same as those described in Study 1.

Results

Data Analysis

Content relevance analysis of the GPSCS.

Screening for discrepant raters. The GPSCS prescreening identified two aberrant judges. One aberrant judge was removed from the content relevance quantitative analysis due to a poor response rate (12 of the 21 quantitative questions were responded to), the qualitative feedback was retained for the subsequent analysis. See Table 3.2 for the content relevance scores of the 16 judges who completed a minimum of 90% of the items, the item means, item medians, and judges JDM scores.

The mean JDM score was 17.63 (*SD* = 7.68) and the median JDM score was 18. Upon inspection of the JDM of the judges, Judge 9 was identified as being considerably greater than other JDM scores[3], therefore this judge was considered aberrant and removed. Thus, the final data set used for the quantitative assessment of item content relevance comprised the responses of 15 judges.

***Quantitative assessment of content relevance ratings*.** Table 3.3 contains the content relevance ratings, mean item scores, median item scores, Aiken's V-coefficients and JDM of the 15 judges. The median relevance rating scores for the 21 items ranged from 2.5 (the midpoint between a fair and good match) to 5 (an excellent match). The median scores for the muscle-tone and muscle-bulk items revealed that all nine items were considered to be a 'very good match' (*Mdn*s = 4). The mean scores for the content relevance of the four muscle-tone items

---

[3] Judge 9 was 2.39 *SD* higher than the mean. The other 7 judges with JDM scores greater than the mean ranged from 0.31 to 0.83 *SD* from the mean score.

ranged from good to very good (*M*s = 3.7 to 4.1). The mean scores for the content relevance of the five muscle-bulk items ranged from good to very good (*M*s = 3.6 to 4.0). The median scores for the muscle-strength items included: one 'excellent match' (*Mdn*s = 5), three 'very good matches' (*Mdn*s = 4), two 'good matches' (*Mdn*s = 3). The mean content relevance score of the muscle-strength items included two good items (*M*s = 3.2 to 3.7) and four very good items (*M*s = 4.0 to 4.3). The median scores for the body-fat items included: five 'very good matches' (*Mdn*s = 4), and one 'fair match' (*Mdn*s = 2.5). The mean content relevance scores for the body-fat items include three very good items (*M*s = 4.0 to 4.1), two good items (*M*s = 3.6 to 3.9), and one fair item (*M* = 2.8).

The V-coefficients of the content relevance scores were analyzed to identify the judges' ratings that were deemed high enough (on average) to claim content relevance. All four of the muscle-tone items, three of the five muscle-bulk items, five of the six muscle-strength items, and four of the six body-fat items had a significant V-coefficient (see Table 3.3). The V-coefficient results identified five items (two muscle-bulk, one muscle-strength and two body-fat items) that were judged to not have high enough content relevance (see Table 3.3). These analyses in combination with the means and median content relevance scores identify items that are suitable for the target construct, and items that may need to be modified or removed from the item set.

Qualitative content relevance assessment of the GPSCS. The qualitative data pertaining to the item-construct matches are listed in Tables 3.4, 3.5, 3.6 and 3.7. According to the qualitative data, the content relevance of muscle-tone was very well supported, as ‘muscle-tone’ was identified by an expert as the most common term used by clients when discussing body ideals. Some experts identified that ‘lack of body-fat’ and ’being toned’ were very difficult constructs to discuss independently among some clients. The content relevance of muscle-bulk was well supported with some concern identified with the terms ‘bulky muscles’ and ‘bulging muscles’. These terms were identified as being very negative or too extreme for some clients. Muscle-strength was generally well supported, although concern arose with the phrase ‘lifting heavy objects’. Fitness experts identified that the exact weight needed be included as part of this item (e.g., able to lift a 100 pounds). The feedback for the body-fat construct included questions and concerns regarding both the straightforward items and the reverse-scored items. Some experts were not sure if the terms ‘skinny’, ‘chubby’ and ‘plump’ were positive or negative attributes. The combination of quantitative and qualitative data will help establish content validity evidence and provide areas that may be modified which will be considered in the discussion.

Table 3.2

Summary of Sixteen Judges' Content Relevance Ratings for the GPSCS

		Judge Identification Number																		
	Item #	1	2	3	4	5	6	7	8	9	10	12	13	14	15	16	17	*M*	*SD*	*Mdn*
Construct																				
Tone	1	5	4	4	4	3	5	4	4	3	4	4	2	3	5	5	4	3.9	.85	4
	2	5	4	4	2	4	3	4	3	3	4	2	3	4	5	5	4	3.7	.95	4
	3	5	3	4	4	4	2	5	3	2	4	3	3	2	4	5	4	3.6	1.03	4
	4	5	5	2	5	4	4	5	4	2	4	4	4	3	5	5	3	4	1.03	4
Bulk	5	5	4	3	3	4	5	5	4	4	4	3	4	2	5	5	4	4	.89	4
	6	5	5	4	5	4	5	3	2	3	3	4	2	3	4	5	2	3.7	1.13	4
	7	5	3	3	4	4	2	5	3	3	3	4	4	4	5	5	3	3.8	.93	4
	8	5	3	1	1	4	5	5	4	1	4	1	4	3	5	5	4	3.4	1.59	4
	9	5	3	4	1	4	2	5	3	2	4	4	3	2	5	5	5	3.6	1.32	4
Strength	10	5	2	5	4	4	5	5	4	4	5	2	4	4	5	5	5	4.3	1.00	4
	11	5	2	5	4	4	5	5	4	1	3	3	4	3	5	5	5	3.9	1.24	4
	12	5	3	4	3	4	2	5	3	3	4	3	3	3	5	5		3.7	.98	3
	13	5	2	2	3	2	2	5		1	3	3	2	2	5	5	4	3.1	1.39	3
	14	5	4	5	2	4	5	5	3	3	4	4	4	3	5	4	4	4	.89	4
	15	5	4	5	2	4	5	5	3	1	4	4	3	4	5	5	4	3.9	1.18	4
Fat	16	5	2	3	3	4	5	5	5	2	4	2	4	4	5	4	5	3.9	1.15	4
	17	5	3	3	3	4	5	5	4	1	4	2	4	2	5	5	5	3.8	1.29	4
	18	5	3	4	3	4	5	5	4	1	4	4	4	3	5	5	4	3.9	1.06	4
	19	5	4	3	5	4	2	5	3	2	4	4	4	4	4	5	5	3.9	1.00	4
	20	5	1	4	2	4	2	5	2	1	3	4	4	3	5	5	5	3.4	1.46	4
	21	2	1	2	2	3	2	5	3	1	3	4	2	2	5		3	2.7	1.23	2
JDM		22	19	16	21	4	24	23	13†	36	7	16	9	17	22	20†	13†			

Note: A total of 21 items are included in the GPSCS. † one missing value, † one value missing, *M* = mean ,Mdn = Median, JDM = Judge's Discrepancy from the Median. The mean JDM score was 17.63 (SD = 7.68) and the median JDM score was 18.

Table 3.3

Summary of Fifteen Judges' Content Relevance Ratings for the GPSCS

		Judge Identification Number																		
	Item #	1	2	3	4	5	6	7	8	10	12	13	14	15	16	17	*M*	*SD*	*Mdn*	V
Construct																				
Tone	1	5	4	4	4	3	5	4	4	4	4	2	3	5	5	4	4.0	.85	4	.75**
	2	5	4	4	2	4	3	4	3	4	2	3	4	5	5	4	3.7	.96	4	.68*
	3	5	3	4	4	4	2	5	3	4	3	3	2	4	5	4	3.7	.98	4	.67*
	4	5	5	2	5	4	4	5	4	4	4	4	3	5	5	4	4.1	.92	4	.80**
Bulk	5	5	4	3	3	4	5	5	4	4	3	4	2	5	5	3	4.0	.93	4	.73**
	6	5	5	4	5	4	5	3	2	3	4	2	3	4	5	2	3.7	1.16	4	.68*
	7	5	3	3	4	4	2	5	3	3	4	4	4	5	5	3	3.8	.94	4	.70*
	8	5	3	1	1	4	5	5	4	4	1	4	3	5	5	4	3.6	1.50	4	.65
	9	5	3	4	1	4	2	5	3	4	4	3	2	5	5	5	3.7	1.29	4	.43
Strength	10	5	2	5	4	4	5	5	4	5	2	4	4	5	5	5	4.3	1.03	5	.82**
	11	5	2	5	4	4	5	5	4	3	3	4	3	5	5	5	4.1	.99	4	.78**
	12	5	3	4	3	4	2	5	3	4	3	3	3	5	5		3.7	.99	3.5	.68*
	13	5	2	2	3	2	2	5		3	3	2	2	5	5	4	3.2	1.31	3	.55
	14	5	4	5	2	4	5	5	3	4	4	4	3	5	4	4	4.0	.88	4	.77**
	15	5	4	5	2	4	5	5	3	4	4	3	4	5	5	4	4.1	.92	4	.78**
Fat	16	5	2	3	3	4	5	5	5	4	2	4	4	5	4	5	4.0	1.07	4	.75**
	17	5	3	3	3	4	5	5	4	4	2	4	2	5	5	5	3.9	1.10	4	.73**
	18	5	3	4	3	4	5	5	4	4	4	4	3	5	5	4	4.1	.74	4	.78**
	19	5	4	3	5	4	2	5	3	4	4	4	4	4	5	5	4.1	.88	4	.77**
	20	5	1	4	2	4	2	5	2	3	4	4	3	5	5	5	3.6	1.35	4	.65
	21	2	1	2	2	3	2	5	3	3	4	2	2	5		3	2.8	1.19	2.5	.45
JDM		21	21	15	23	4	24	21	14†	5	17	11	19	20	18.5†	11.5†				

Note: The removal of the 2 aberrant judges responses resulted in the following. The mean JDM score was 16.33 (SD = 6.17) and the median JDM score was18.5.
† one value missing, *M* = mean , *Mdn* = Median, V = Aiken's (1985) item content validity coefficient.* $p < .05$,** $p < .01$.

Content Representativeness Analyses of the GPSCS.

Screening for discrepant raters. The first step in the content representativeness analyses was to screen for discrepant raters using the same methods as in the content relevance analyses. One judge (judge 9) was identified as a discrepant rater due to the considerably greater JDM scores than the group of judges, therefore this judge was removed from the analysis; see Table 3.8.

Item representativeness. Table 3.9 displays the judges' ratings for the muscle-tone, muscle-bulk, muscle-strength and body-fat item sets as well as the descriptive statistics associated with each rating. The mean content representativeness ratings for the item-sets were as follows; muscle-tone 4.13 (SD=.81), muscle-bulk 3.63 (SD=1.09), muscle-strength 3.88 (SD=.72) and body-fat 3.73 (.88). The mean and median ratings for the four item sets ranged from "good" (3) to "very good" (4) which is an indication that the fitness experts felt that the item sets adequately covered the construct of interest.

The Aiken's V-coefficient (Aiken, 1985) was used to statistically analyze the content representativeness ratings of the item sets according to the construct of interest. A V-coefficient value close to 1.0 indicates "excellent" content representativeness and a score of 0 indicates "poor" content representativeness. As seen in Table 3.9, the V-coefficient for muscle-tone was .78 ($p < .01$), muscle-bulk was .66 ($p < .05$), muscle-strength was .72 ($p < .01$), and body-fat .68 ($p < .05$). These coefficients provide statistical support for the content representativeness of all four item sets.

Table 3.4

Content Relevance and Representativeness of the Muscle-tone Items

Content Relevance	
Item	Comments
1. My muscles appear toned.	- appearing toned to me is absence of subcutaneous fat, appearing lean - "toned" is mentioned often by primarily female clients and in the context of the above operational definition - best match - doesn't describe toned - toned is the most common used term from clients in my experience.
	- again, absences of fat may cause appearance of firm muscle or firm can be related to active muscle - perhaps an example of firm? however this may change their current self perception - may be confused with muscle-bulk - I like toned better than firm - firm doesn't capture the essence of all the definition - a lot of people see physical "jiggling" as the opposite of toned
3. My muscles appear lean.	- if lean means absence of fat rather than lean muscle tissue as in small/spindly - not a term used in conversations with clients - lean could be viewed as skinny or no muscle-bulk which could be confusing for general public; could be seen as related to muscle-bulk - I don't think the general population would use the word lean or know what it looks like. I think of chicken wings. - - great for an athletic audience however in a sedentary population individuals may not understand exactly what you mean by lean.
4. My muscles appear well defined.	- clients use the word toned to mean smaller and defined - I am not sure women associate definition with tone as much as muscle-bulk - well defined muscles appear toned muscles can be well defined or absence of fat, not necessarily bulky - when clients state toned they are looking for the "well defined" image - I think women would become concerned with "too well defined. I think long and lean would be a good way to describe this. - very popular statement in Personal training field. - I wonder if people might have an issue with defined... because it’s all likely to fall into comparisons

(continued)

Content Representativeness Comments and Suggested Terms	
Comments:	- good except kinesthetic feel - well defined - not as specific - on their own can be confusing to the general public, firm and lean are not good representations by themselves but when clumped together with other words such as defined they help describe tone - I believe item 1 word toned is not a representation. the individual should define toned with stuff like other items, firm etc. - I think all items represent muscle-tone very well
Other Terms/Words:	- lean, long muscles, don't want bulk, females ask for muscle toning exercises - use the word firm. my clients are all about appearance, they use the term definition but don't really get that that body-fatness is a huge part of tone - long lengthened - cut ripped - being ripped - long, ripped, cut, juiced, sexy - tight - mostly toned, definition, strength - may be sculpted, often discussed with removing excess fat/flab - long and lean, smooth lines, then when you flex they pop out, avoid big - hard muscles? - not bulky, defined, I can see the shape of the muscle. - I think you got it!

Table 3.5

Content Relevance and Representativeness of the Muscle-bulk Items

Content Relevance	
Item	Comments
5. I appear to have large muscles.	- perhaps relates to strength more than bulk - again could have overlying adipose that makes muscles appear larger than they are - large versus huge - kind of vague maybe have sub questions comparing to the general public or to the persons fitness levels through out their life ie)this is the largest I have ever been - doesn't capture the whole essence - large is good for boys, bulky is too big - large muscle bilk, great match
6. My muscles appear bulky.	- bulky is a negative word - fat could be bulky. how does someone define bulky? some may think about muscle but ,ay be more around fat? - negative term=bulky - maybe be more specific to what large and bulky mean - using bulky does describe what clients want with bilk, bigger thicker, more dense etc. - if this is something we are looking to reduce to improve functionality
7. My muscles appear to be huge.	- depends on past experiences, "huge" is often used with men but not as much women - I think hug relates to strength more than bulk - exaggerated term, perhaps not used as often, sounds conceited?? - someone might have fat around there muscle and think the muscle appears huge - huge doesn’t capture the look of it. yes, big but also need defined - some will want this most don't - I think bulky and huge are a great match .
8. My muscles appear small.	- small is more toned - small = not bulky or large - I like this one the best. even skinny people don't want small muscles thus need to increase bulk a bit - I am scrawny
9. I appear to have bulging muscles.	- non-educated athletic people seem to only comment on bulging muscles on "lean". so i think it is less about bulk and size - small muscles can also be bulging if active and in absence of fat - exaggerated term not as preferable - maybe be more specific on what makes a muscle bulging opposed to large or huge. or describe differences - seems to capture more of the essence - I guess it all depends on where the bulges are... - like this

(continued)

Content Representativeness Comments and Suggested Terms	
Comments:	- item 9 is redundant - bulging doesn't=large / bulky - bulging muscles - seems a bit exaggerated compared to the rest - don't like bulging or bulky. many of my clients refer to overweight people as bulging or bulky. - need to be muscle mass, which could be relevant to skeletal size - have examples for each i.e. bulky body builder, small but fit long distance runner? - I think I would leave out item 6 and 9 as bulky or bulging are too open and the client would feel more comfortable staying their body structure as large, huge etc. - I see the items above having a negative connotation. maybe bulk needs to be changed - Bulky is still a struggle
Other Terms/Words:	- guys want beach type body which is big biceps shoulders and arms - bulk strictly refers to size with my clients. mostly they talk about something not fitting the same way it used to. all females view size increases as negative. bulk=bigger but not necessarily BIG - big, large - BIG - beefy - appear to have more muscle-bulk than fat? - massive, pumped, jacked - muscle growth, larger muscles, when I talk about bulking up with a client we start with a base for comparisons usually girth measurement - ripped, sculpted, massive, build muscle, more defined - large huge small - range of motion, functional mass versus bulk, big beach body - thick - bigger, gains weight/gains mass - using talk about gaining size. that's pretty general though.

Table 3.6

Content Relevance and Representativeness of the Muscle-strength Items

Content Relevance	
item	Comments
10. I am physically strong.	- maybe give an example or get the person to think of what makes them strong or not strong - - context could be taken to mean strength other than muscle
11. I am physically weak.	- people under estimate themselves with strength - - somehow I think about health and illness with this statement
12. I am strong enough to lift heavy objects.	- heavy is relative and perceived differently across gender and age categories. however, the question gives an idea of where the person feels they are t - "heavy objects" is subjective likely based on how strong the one saying it is - examples? - what is really heavy? barbells, Olympic plates or moving furniture? Perhaps people will have difference in opinion of what is heavy? - how heavy? - catch 22. strength is relative i.e. what is heavy? - heavy objects for me or the general public - a very ### need statement, great for any audience - heavy objects aren't always about strength. do they have power.
13. I am physically feeble.	- feeble is more of a character of old age and bone strength than the general population of strength - will people know what feeble means? examples? - weak better more commonly used term than feeble - change feeble to weak - what's feeble? - negative - feeble needs to be described for the client - feeble creates the image of aged to me - great link to muscular strength
14. My muscles are strong.	- muscles may be strong but may not know how to use them to lift heavy objects - again what is this relative too? isokinetic, concentric, eccentric strength, maximal strength? - would you compare this with your most fit time, or against the general population or no comparison and just how you feel at that point in time - strong muscles great match
15. My muscles are weak.	- muscles may be weak, but know how to use them to lift heavy objects - some clients may have strong lower bodies and really weak upper bodies? - again what is this relative too? isokinetic, concentric, eccentric strength, maximal strength? - this is less the whole person, muscles can change

(continued)

Content Representativeness Comments and Suggested Terms
Comments: - I don't like 13 it refers to overall function and old age - not sure about feeble - would a comment like I am not strong enough? or I am too weak to __work? - perhaps define strength (e.g. the ability to produce force, e.g. one repetition max) - I am physically strong/weak may be misunderstood. One individual could be physically strong muscle wise but see themselves as physically weak due to a possible injury... - strength is relative also different types of strength isokinetic, concentric, eccentric - feeble versus weak - I think all of the items here rightly describe for the clients the perception they have of their muscle-strength - I read 13 as a health vs. physical measure. :my body is physically weak" - great representation, all items definitely cover muscle-strength - feeble...not sure a guy would admit to being feeble Muscle-strength terms/words: - strength has to do with being fit and being able to do things more easily - anytime they talk about strength its relative to an exercise or a weight used, and therefore less about overall strength. they usually use vague terms like "weak" or "strong" but size and firmness - functional - endurance, power - powerful, tough - I can lift more than __lbs? I like my muscles are physically strong /weak. Powerful. - rate of force development strength speed, speed strength, isokinetic , eccentric concentric - pump iron, press x amount of weight, muscles are pumped - stronger, ability to lift or hold - mass to weight ratio (I can lift lots for my size) functional strength (I can physically do what I want to) endurance vs. power. - opening a really heavy door

Table 3.7

Content relevance and representativeness of the body-fat items

Content Relevance	
item	Comments
16. My body appears fat.	- I am fat? - appear fat can be being larger than a peer maybe more muscled - how hard is it for people to say they are fat? - excellent statement - I am too fat, I have too much fat on my body
17. My body appears skinny.	- if skinny=small amount of fat skinny can mean small muscles (lean thin muscles) with large proportion of fat - still looks small - people’s idea of skinny is not consistent. could be a good thing. - it is negative to be skinny too - compared to norms?
18. My body appears thin.	- can be related to spindly muscles not necessarily body-fat. less extreme version of item 17 - compared to norms? - I think this is more desirable than skinny in general public - body appears slim could be another option
19. My body appears lean.	- both muscle-tone and body-fat, people want to be lean not skinny - more tone? - lean=absence of fat - average person uses fat or skinny/thin. people more aware of actual body composition will use "lean" - I like this the best - great match, but sounds like a statement targeting more of an athletic population as oppose to general population (they may interpret lean as skinny)
20. I appear chubby.	- chubby most often thought of as fat could also appear chubby but have large muscles - a pejorative term - people may be resistant to use term over "body appear fat" - using "I" versus "my body" I may be better than my body for self-perceptions - I don't like the word chubby. what does chubby mean? - negative word - what is the difference or describe the difference between chubby and fat and plump - this will get a lot of "yes" responses I think it insinuates a bit extra but not at high risk - I'm jiggly. My body jiggles when I move. fat kid in grade 4? - overweight perhaps?
21. I appear to have a plump body.	- plump is not a common word, something older people would say - not common language - plump same as item 20 seems to mean more/larger than chubby - plump also pejorative - other terms used by my clients = flabby, slightly overweight - is plump good or bad? - negative word - what is the difference or describe the difference between chubby and fat and plump - seems open to definition with a description of plump - plump makes me think of women in the Victorian times.

(continued)

Content Representativeness Comments and Suggested Terms
Comments: - I am not sure what you are using these terms for but it is not often that people use these terms to refer to their whole body, most of the time it is specific to body parts. - 21 is redundant - absence of "tone" on a skinny body doesn't mean absence of fat a "skinny" body appears to have fat if not toned - don't like plump or chubby - apple, pear all over distribution of body-fat - I think I would leave item 21 out. the other items define body-fat nicely - I like the use of plump and chubby. to me it implies more than I want versus statistical norms - items are very understandable and convey the construct very well - flabby, soft
Body-fat term/words: - extra tire around the middle, women usually just point to areas that they want less fat inner thigh, arms, tummy - never heard of "chubby" or "plump". I hear the terms flabby, soft or excess - obese, overweight, excess, over fat - chunky - overweight, obese - flabby, overweight - overweight, obese, BMI, underweight - blubba, lard, non-functional mass - location specific - BMI terminology - flab - central vs. peripheral fat, cosmetic fat loss versus health concerns, body types can influence body shape, we can improve what you have - overweight, slim, slender, tiny - - lose weight

Table 3.8

Item-Set Content Representativeness Ratings and Descriptive Statistics of the GPSCS of the Seventeen Judges.

	Judges																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	*M*	Mdn	*SD*	R
Muscle-tone	5	3	4	5	4	5	5	4	3	4	3	4	3	3	5	5	4	4.06	4	.83	3
Muscle-bulk	5	3	3	3	4	5	5	3	1	4	2	3	3	2	5	5	3	3.47	3	1.23	5
Muscle-strength	5	4	4	4	4	5	4	3	1	4	3	3	4	3	5	4	3	3.71	4	.99	5
Body-fat	5	3	4	2		5	4	3	4	3	4	3	3	4	5	4	4	3.75	4	.86	4
JDM	5	2	0	3	1†	5	3	2	6	2	3	2	2	3	5	3	1				

Note. †not all constructs were responded to therefore median was used as the score in the assessment of JDM.
The mean JDM score was 2.82 (*SD* = 1.63) and the median JDM score was 3.

Table 3.9

Item-Set Content Representativeness Ratings and Descriptive Statistics of the GPSCS of the Sixteen Judges

	Judges																				
	1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	*M*	*Mdn*	*SD*	R	V
Muscle-tone	5	3	4	5	4	5	5	4	4	3	4	3	3	5	5	4	4.13	4	.81	3	.78**
Muscle-bulk	5	3	3	3	4	5	5	3	4	2	3	3	2	5	5	3	3.63	3	1.09	4	.66*
Muscle-strength	5	4	4	4	4	5	4	3	4	3	3	4	3	5	4	3	3.88	4	.72	3	.72**
Body-fat	5	3	4	2		5	4	3	3	4	3	3	4	5	4	4	3.73	4	.88	4	.68*
JDM	5	2	0	3	1†	5	3	2	2	3	2	2	3	5	3	1					

Note. †not all constructs were responded to therefore median was used as the score in the assessment of JDM.
The mean JDM score was 2.63 (*SD* = 1.45) and the median JDM score was 2.5.
*= p < .05,**= p <. 01.

Qualitative Assessment of Content Representativeness Ratings. Judges commented on the representativeness of the group of items relative to the constructs of interest. Comments were made about specific items within the group that could be modified to improve the overall representativeness of the group of items. Tables 3.4, 3.5, 3.6 and 3.7, include the comments about the representativeness, overall comments and identification of terms used between fitness professionals and clients to discuss these constructs. The qualitative feedback associated with the content representativeness of the muscle-tone, muscle-bulk, muscle-strength and body-fat item-set revealed that overall the items were well endorsed. Fitness experts also mentioned that clients often discuss their body by simply pointing to a specific area of concern rather than talking about it.

Muscle-tone. Participants provided a list of terms used in their discussions with clients about the muscle-tone of their body. They indicated that the terms used included: lean and long muscles, firm muscles, cut and ripped muscles, juicy muscles, tight and sexy muscles, and sculpted muscles (see Table 3.4 for the complete set of terms).

The fitness experts identified that item 3, (My muscles appear lean) was a confusing item (see Table 3.4). Experts suggest that lean muscle is not a term used in conversation and more sensible when referring to meat (i.e., chicken). One expert suggested the term may be understood by athletic populations but not the general population, therefore despite a good mean and median content relevance score and a significant V-coefficient this item was removed from the item set. The fitness experts provided many suggestions for creative descriptors of muscle-tone which included “being ripped” and “sculpted”. These two descriptors of muscle-tone were developed into the following items that will be tested in the next study: “My muscles are ripped” and “My muscles are well sculpted”. These two new items will be included as measures of muscle-tone.

Muscle-bulk. Terms used to describe muscle-bulk in discussions between fitness professionals and their clients about the body included: beach body, size, big, beefy, massive, pumped, jacked, ripped, sculpted, and thick (see Table 3.5). Feedback from the fitness experts suggested that, typically, weightlifters or people wanting to put on some muscle mass say they want to “get big”. Ridgeway and Tylka (2005) found that some groups of men idealize having big biceps, a thick neck, large and broad shoulders, thicker and defined legs, “a chiseled six pack”. Many of the descriptors identified by Ridgeway and Tylka (2005) provided further support for the desire to have “big” muscles. Furthermore, Leit, Gray and Pope (2002) found that society’s view of body ideals for men among men was the bigger the muscles the better. These ideas were addressed by constructing a new item “My muscles appear big” which will be used in the next validity study.

Although these muscle-bulk terms were popular among some clients the fitness experts suggested that these body ideals were not popular among all fitness participants. The muscle-bulk feedback identified that some terms describing muscle-bulk may have negative connotations with some populations (i.e., women) or may be too extreme for the average person to endorse (e.g., bulging muscles, bulky, huge muscles). These comments were examined in accordance with the operational definition and specified items. Secondly, one item ‘I am physically feeble' was identified by a few fitness experts as being beyond the operational definition of muscle-bulk as well as being too negative to include in this set of items. They therefore suggested it be removed.

In summary, as noted in Table 3.10, 12 items were retained, 9 items were removed and 5 items were added. The GPSCS included 17 items for Study 3.

Muscle-strength. Terms used in client-fitness professional discussions to express muscle-strength included: being powerful, tough, able to lift more than 'X' amount of weight (see Table 3.6 for more information). None of these terms seemed appropriate in view of the study with the academic experts therefore the item-pool was left the same (the items are listed in Table 3.10).

Body-fat. Body-fat terms used in client-fitness professional discussions about the body include: extra tire around the middle, 'they just point', blubba, overweight, flab, etc. (See Table 3.7 for more detail). Two items were added to the body-fat construct ("My body jiggles" and "my body is flabby"). The three reverse scored items were removed. The body-fatness items listed in Table 3.10 will be assessed in Study 3.

Instrument Developed from the Content Relevance and Representativeness of the GPSCS. The same procedures from Study 1 were used to develop the items to be included in the GPSCS for the next two validity studies. See Table 3.10 for the items that were generated as a function of this study.

Table 3.10

GPSCS Items Included/ Removed/ Modified From Study 2 For Study 3

	GPSCS Item	
Construct	Items Reviewed in Study 2	Items Developed for Study 3
Muscle-tone	1. My muscles appear toned.	1. My muscles appear toned. (retained)
	2. My muscles appear firm.	5. My muscles appear firm. (retained)
	3. My muscles appear lean.	(removed)
	4. My muscles appear well defined.	11. My muscles appear well-defined. (retained)
		14. My muscles are ripped. (new)
		17. My muscles are well sculpted. (new)
Muscle-bulk	5. I appear to have large muscles.	2. I appear to have large muscles. (retained)
	6. My muscles appear bulky.	8. My muscles appear bulky. (retained)
	7. My muscles appear to be huge	6. My muscles appear to be huge. (retained)
	8. My muscles appear small.	(removed)
	9. I appear to have bulging muscles.	(removed)
		15. My muscles appear big. (new)
Muscle-strength	10. I am physically strong.	3. I am physically strong. (retained)
	11. I am physically weak.	7. I am physically weak. (retained)
	12. I am strong enough to lift heavy objects.	(removed)

(continued)

	GPSCS Item	
Construct	Items Reviewed in Study 2	Items Developed for Study 3
	13. I am physically feeble.	(removed)
	14. My muscles are strong.	9. My muscles are strong. (retained)
	15. My muscles are weak.	12. My muscles are weak. (retained)
Body-fat	16. My body appears fat.	4. My body appears fat. (retained)
	17. My body appears skinny.	(removed)
	18. My body appears thin.	(removed)
	19. I body appears lean.	(removed)
	20. I appear chubby.	10. I appear chubby. (retained)
	21. I appear to have a plump body.	(removed)
		13. My body jiggles. (new)
		16. My body is flabby.(new)

Note. A total of 17 items will be included in the GPSCS in Study 3. Retained items (n = 12) have identical wording in Study 2 and Study 3. Removed items (n = 9) were removed after Study 2 and will not be included in Study 3. New items (n = 5)were not included in Study 2 and will be included in Study 3.

Discussion

The purpose of this study was to examine the content relevance and content representativeness of muscle-tone, muscle-bulk, muscle-strength and body-fat items according to an expert review process conducted by fitness professionals. The item content of the GPSCS was previously adapted according to results and feedback from the content validity study that used an academic expert panel (in Study 1). The current study found the items demonstrated good relevance and representativeness according to the constructs of interest. The feedback and results from the fitness professionals were used to further modify the instrument prior to assessing the factorial validity.

Muscle-tone

The content relevance and representativeness of the muscle-tone items were good to very good and had significant *V*-coefficients. According to the fitness experts' feedback, personal training clients and fitness class participants most often look for help in achieving a toned body, like those seen in popular fitness magazines (Labre, 2005; Law & Labre, 2002; Leit et al., 2002). The fitness experts suggest a toned body is required to be lean and muscular which increases the complexity of muscle-tone. According to the fitness professionals, some personal training and fitness class participants have difficulty recognizing the difference between a lean body with little fat and a toned body. For example, some people have very low levels of body-fat which allows their small muscles to be visible, while others have muscle development and a low level of body-fat. Furthermore, experts identified that an overlap between the body-fat and muscle-tone constructs should be expected, as for one's muscle-tone to be visble, one must have low levels of body-fat. These comments provided support for the operational definition of muscle-tone as well as confirmed the importance of maintaining conceptual distinction between muscle-tone and body-fat.

The fitness experts judged Item 3 "My muscles appear lean" as having good content relevance and a significant V-coefficient. Although the quantitative results supported the retention of Item 3, the qualitative feedback included many comments opposed to the inclusion of the item in the inventory. The term lean muscle was considered confusing, especially for non-athletic populations and possibly more appropriate for the discussion of poultry than people. One expert suggested that an athletic population (such as the one completing this content relevance instrument) would understand the term while the general population may be confused or unsure. Another comment suggested that the terms lean and firm are usually included together to clearly define muscle-tone. Furthermore, the experts provided multiple common terms that describe muscle-tone such as "well-defined" and "ripped" that were used regularly in conversations with clients. These terms were used in two new muscle-tone items.

Muscle-bulk

The muscle-bulk items demonstrated good relevance to the construct of interest however two items were flagged as being potentially problematic (Items 8 and 9). The median relevance score for all muscle-bulk items was very good (*Mdn* = 4). The validity coefficients of three of the five items were significant, and feedback indicated that some terms used may have a negative tone or exaggerated descriptors.

The first item that received some scrutiny was “My muscles appear small” (Item 8) as this item did not reach a significant validity coefficient and the qualitative feedback suggested a better item fit for the muscle-tone construct than the muscle-bulk construct. On the contrary, feedback was also received stating this item is a good fit “to represent not being bulky” or “not having large” muscles. Researchers have found that including reverse scored and positively worded items in the same construct may result in poor performance of the construct (Currey, Callahan, & DeVellis, 2002). Due to inconsistent support for the item in the qualitative and quantitative analyses of this study, and the conflicting recommendations for including both positively worded and reverse scored items in a scale (DeVellis, 2012), the item was removed from the muscle-bulk item-set.

A second item “I appear to have bulging muscles” (Item 9) was considered problematic. The median relevance score for this item was very good (*Mdn* = 4) and mean score was good (*M* = 3.7). However, the validity coefficient was not significant (V = .43) and feedback that “bulging” refers to the shape of the muscle rather than the size of the muscle brought about further reservations. The fitness experts suggested that one may have good muscle-tone and define their muscle as bulging without having bulky muscles. A second set of feedback suggested that the term ‘bulge’ may be confusing to respondents as ‘bulge’ can be used in reference to muscle bulges and fat bulges. Although the term in the item is specifically 'muscle bulge' the fitness experts suggested that respondents may be distracted by the word 'bulge' and focus more on the body-fat bulges. Due to the potential confusion around the term bulging expressed in the qualitative data and the quantitative results, "I appear to have bulging muscles" (Item 9) was removed from the item set for muscle-bulk.

Some terms identified by the fitness experts for muscle-bulk and muscle-tone were identical. For example, fitness experts identified the terms "being ripped" and "sculpted" as muscle-tone items; the terms "ripped" and "sculpted" were also identified as muscle-bulk items. Self-perception of muscle-tone and muscle-bulk constructs were developed as separate constructs; the experts agreed that these are separate constructs, although experts used similar terms in defining the constructs. The crossover in descriptors used for muscle-tone and muscle-bulk may be due to issues related to construct specificity and clarity. Definitions of muscle-tone and muscle-bulk may vary according to the situation or the experience and characteristics of the sample. The possibility of muscle-tone and muscle-bulk crossing-over or tapping into related constructs can be problematic (DeVellis, 2012).These terms were incorporated into the GPSCS and were examined closely in the following studies.

Muscle-strength

Six muscle-strength items were assessed. The results indicated that four of the six items (Items 10, 11, 14 and 15) had a median rating indicating very good (*Mdn* = 4) to excellent (*Mdn* = 5) content relevance, a mean content relevance rating indicating a very good score (*M* = 4.0 to 4.3), a significant validity coefficient, and positive comments supporting the use of the items to represent the construct of muscle-strength.

Two muscle-strength items were highlighted as problematic. The first item of concern was the item “I am physically feeble” (Item 13). The median and mean content relevance scores were good, however the validity coefficient was non-significant and the comments did not support keeping this item. The fitness expert comments suggested that ‘feeble’ suggests old age, poor bone strength and density, and would be considered a derogatory label. The fitness experts’ interpretations were not those intended or desired. Given this concern, this item was dropped from the item set.

A second problematic item was “I am strong enough to lift heavy objects” (Item 12). The median and mean content relevance score for the item was good and the validity coefficient was statistically significant. Despite these supportive content-relevance ratings, many of the fitness experts were concerned with the item ambiguity surrounding ‘what constitutes a heavy object’ and ‘what a heavy object is not’. The fitness experts were concerned about the inconsistency among responses to this item. For example, one may consider her/himself able to lift heavy objects because s/he can bench press their body weight, while others believe they can lift heavy objects because they were able to lift up a four-litre jug of milk. This item was also deemed as problematic among the academic experts. Due to all these concerns, the item was dropped from the item set.

Body-fat

Five of the six body-fat items had a median content relevance score of very good, three of the body-fat items had a mean content relevance score of very good and the other three items had a mean content relevance score of good, and four of the items had a significant validity coefficient. The experts endorsed many of the items and suggested revisions for others. Three items (Items 17, 20 and 21) were identified as problematic according to the quantitative results and the qualitative feedback.

According to the quantitative data, the item “I appear chubby” (Item 20) had good to very good relevance (*M* = 3.6, *Mdn* = 4), while the non-significant validity coefficient and inconsistent feedback suggested that some experts did not fully endorse this item. The fitness experts offered a variety of comments including questioning “what does chubby look like?”, suggesting that chubby is a “pejorative term” and even positive comments such as “I like this one, it insinuates a bit extra”. Due to the inconsistency among the experts’ interpretations, this item was kept for further consideration in the next study.

The item “I appear to have a plump body” (Item 21) was deemed as problematic with fair content relevance (*M* = 2.5, *Mdn* = 2.8), a non-significant validity coefficient and unfavourable content relevance feedback. One fitness expert stated that “plump makes me think of women in the Victorian times”. The term plump is not commonly used among young people, therefore not a part of the young adult micro-culture. Due to the obsolete nature of the term, respondents may not be able to identify with this statement. This item was subsequently removed from the item set.

Conflicting results were obtained for Item 17 (my body appears skinny). Good to very good content relevance (*M* = 3.9, *Mdn* = 4.0) and a significant validity coefficient were observed (see Table 3.3). Despite, the positive quantitative results, the feedback from the fitness experts suggested that “skinny” refers to both a small amount of fat and a small amount of muscle which makes this item ambiguous. Secondly, being skinny is considered a negative trait which is not parallel to being ‘not fat’, as skinny suggests very low body-fat and very low muscle-tone. Skinny was one of the words used by male University students to describe the body they do not desire (Ridgeway & Tylka, 2005). This item was removed due to the qualitative feedback concerning this item.

Body-fat is a common topic of conversation among fitness experts and their clients which may explain the numerous “fat” terms identified by the fitness experts. The most common terms used by the professionals and clients included “flabby” and “jiggly” (see Appendix D for a full list of the terms identified by the experts). In order to reflect and use the language that the population of interest uses to talk about body-fat, the items “my body is flabby” and “my body jiggles” were added to the item set for the next validity study.

Overall, the fitness experts indicated very good content relevance and supported the majority of the items for the GPSCS. As a result of Study 2, nine items were removed from the GPSCS (Items 3, 8, 9, 12, 13, 17, 18, 19 and 21), five new items were added and 12 items (Items 1, 2, 4, 5, 6, 7, 10, 11, 14,15, 16 and 20) were retained from Study 2. The fitness experts' input into the GPSCS was a valuable step in distinguishing current physical self-concept terminology. Modifications were made to the proposed item content of the GPSCS and initial validity evidence supporting the GPSCS was gathered. As viewed in Table 3.10, the item numbers have been modified to represent the item numbers in the GPSCS instrument package and these item numbers will remain the same for Studies 3 and 4.

References

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*, 131-142. doi:10 .1 177/0013164485451012

Currey, S. S., Callahan, L. F., & DeVellis, R. F. (2002). *Five-item Rheumatology Attitudes Index (RAI): Disadvantages of a single positively worded item* (Unpublished paper). University of North Carolina, Chapel Hill, NC.

DeVellis, R. F. (2012). *Scale development: Theory and application (3rd ed.)*. London, GB: Sage.

Dunn, J. G. H., Bouffard, M., & Rogers, W. T. (1999). Assessing item content-relevance in sport psychology scale construction research: Issues and recommendations. *Measurement in Physical Education and Exercise Science, 3*, 15-36. doi:10.1207/s15327841mpee0301_2

Gotwals, J. (2006). *Developing a perfectionism measure in sport* (Unpublished doctoral dissertation). University of Alberta, Edmonton, AB.

Hellsten, L. M. (2005). *The development and validation of a time management scale for exercise adoption, participation, and adherence* (Unpublished doctoral dissertation). University of Alberta, Edmonton, AB.

Labre, M. P. (2005). The male body ideal: Perspectives of readers and non readers of fitness magazines. *The Journal of Men's Health and Gender, 2*, 223-229. doi:10.1016/j.jmhg.2005.03.001

Law, C., & Labre, M. P. (2002). Cultural standards of attractiveness: A thirty-year look at changes in male images in magazines. *Journalism and Mass Communication Quarterly, 79*(3), 697-711. Retrieved from http://www.aejmc.org/home/publications/jmc-quarterly/

Leit, R. A., Gray, J. J., & Pope, H. G. (2002). The media's representation of the ideal male body: A cause for muscle dysmorphia? *International Journal of Eating Disorders, 31*, 334-338. doi: 10.1002/eat.10019

Leit, R. A., Pope, H. G., Katz, D. L., & Oliva, P. (2001). Cultural expectations of muscularity in men: The evolution of Playgirl centerfolds. *International Journal of Eating Disorders, 29*, 90-93. doi: 10.1002/1098- 108X(200101)29:1<90::AID-EAT15>3.0.CO;2-F

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

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performance as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749. doi:10.1037// 00 03-066X.50.9.741

Ridegeway, R. T., & Tylka, T. L. (2005). College men's perceptions of ideal body composition and shape. *Psychology of Men and Masculinity, 6*, 209-220.

Study 3

Establishing Construct Validity Evidence for the Internal Structure of the General Physical Self-Concept Scale in University Students

The physical self is an important aspect of the self-system; appearance, physical attributes and physical abilities of the body function as a substantial interface between the person and the world (Fox, 2000). The body plays a major role in social communication (Jackson, Hunter, & Hodge, 1995), mental health (Feingold, 1992), confidence in social interactions (Eagly, Ashmore, Makhijani, & Longo, 1991; Nezlek, 1999), perceived intelligence (Jackson, Hunter, & Hodge, 1995), relationships (Peretti & Abplanalp, 2004), and perceived competence (Parks & Kennedy, 2007). A positive physical self-concept has been identified as an important aspect of socialization whether these social phenomena are conscious or subconscious.

Individual differences, on the basis of gender, culture, ethnicity and sexual orientation account for variations in body ideals (Thompson, Heinberg, Altabe & Tantleff-Dunn, 2002). Among North American women, an attractive female body can be described as slender, having little body-fat, cellulite-free, having some muscle definition and having a small waist-to-hip ratio to create an hour glass figure (Dixon, Grimshaw, Linklater, & Dixson, 2011; Thompson, Heinberg, Altabe, & Tantleff-Dunn, 2002). Among North American men, an attractive male body includes a toned chest, sculpted ‘six-pack’ abdominals, big biceps and broad shoulders (Ridgeway & Tylka, 2005). These body ideals include perceptions of body-fat, body size, body shape and muscularity as the foundations for evaluating the physical self.

Although muscularity at least contributes to the factors used to describe and self-evaluate physical self-concept, there is no existing measurement instrument that assesses physical self-concepts of muscle-tone and muscle-bulk. The current study was the third in a series of four studies focused on providing validity evidence to support the use of the GPSCS to measure physical self-concepts of muscle-tone and muscle-bulk. Initial validity evidence supporting the items in the GPSCS was established in two studies that assessed the content relevance and content representativeness of the items designed to measure these constructs with academics and fitness experts. The purpose of this paper was to develop further validity evidence supporting the use of the GPSCS as a measure of muscle-tone, muscle-bulk, muscle-strength and body-fat.

Construct validity is a concept that unifies content, substantive, structural, generalizability, external and consequential aspects of validity (Messick, 1995). This study specifically focused on the structural aspects of validity by examining the internal structure of muscle-tone, muscle-bulk, muscle-strength and body-fat items that are contained within the GPSCS. Structural validity evidence, when considered in combination with the theoretical understanding of the constructs and content validity evidence, provides a unified interpretation of findings.

In this study, the internal structural validity of the GPSCS was examined using exploratory factor analysis (EFA). More specifically, EFA was used to determine the latent structure underlying the GPSCS items that were intended to measure muscle-tone, muscle-bulk, muscle-strength and body-fat among a sample of university students.

Method

Participants

A total of 352 (155 males and 197 females) undergraduate health education students from two Canadian Universities participated in the study. The students volunteered to participate in the study and did not receive class credit for participation. The students had a mean age of 20.5 years (*SD* = 3.12). The body mass index (BMI) calculated from self-reported height and weight was 24.2 kg/m^2 (SD =7.18). The mean leisure time activity level score on the Godin Leisure Time Exercise Questionnaire (GLTEQ) was 72.90 METs per week (54.49). According to Garcia Bengoechea, Spence and McGannon (2005), achieving 38 METS per week for men and 35 METS per week for women is a sufficient level of physical active.

Instruments

Demographics. The demographic questions asked participants to indicate their sex, age, height and weight (See Appendix H for the complete Study 3 survey). Self-report height and weight were used to calculate BMI.

Godin Leisure-Time Exercise Questionnaire (GLTEQ). The GLTEQ assessed self-reported leisure-time physical activity (Godin & Shephard, 1997) (See Appendix H for the complete Study 3 survey). Participants consider an average 7-day period and recalled the number of times they participate in 15 minutes of strenuous, moderate or mild exercise during their free-time. Strenuous exercise was described as exercise that results in the heart beating rapidly while doing activities such as running, hockey, soccer, squash, etc. Moderate exercise was described as exercise that was not exhausting and includes exercise such as walking, tennis, easy swimming, etc. Mild exercise was considered exercise that requires a minimal effort, such as yoga, bowling, golf, etc. The frequency of strenuous, moderate and mild exercise was multiplied by 9, 5, and 3 METs respectively and then summed (Godin & Shepard, 1997). The GLTEQ was used to describe the physical activity level of the sample. According to Garcia-Bengoechea, Spence and McGannon (2005), men expending 38 METs or more per week and women expending 35 METs or more per week were considered sufficiently active to achieve health benefits.

General physical self-concept scale (GPSCS). The GPSC contains 17 items proposed to measure four constructs: *muscle-tone* (5-items), *muscle-bulk* (4-items), *muscle-strength* (4-items), and *body-fat* (4-items). The operational definition for each construct is listed below:

Muscle-tone: The self-perception of the degree to which muscles appear lean, firm and defined.

Muscle-bulk: The self-perception of the degree to which muscles appear large and bulky in size.

Muscle-strength: The self-perception of the presence or absence of muscle strength and the ability to lift heavy objects.

Body-fat: The self-perception of the degree of fat that appears on the body.

Participants were instructed to indicate on a scale ranging from (1) *strongly disagree*, to (5) *strongly agree* the extent to which they agreed or disagreed with the declarative statements about their body (See Appendix H for the complete Study 3 survey package). Two muscle-strength items were negatively worded (Item 7 and Item 14) and were reverse scored prior to data analysis.

Procedures

Following ethical approval from the institutional research ethics boards, students were recruited to volunteer in the study. Participants were informed both orally and in writing that participation was voluntary and they were free to withdraw without consequence at any time. Implied consent was provided by the students if they completed and returned the survey package. The survey package was distributed and returned in a classroom setting. The complete survey package took 15-minutes or less to complete.

Data treatment. The data were subjected to principal axes (PA) factor analysis followed by a direct oblimin transformation (delta=0) using the PASW Statistics 18. The cases were excluded pairwise which indicates that the missing data points were excluded if the data were required for the specific analysis.

Two vital components of sampling procedures for EFA were sample size and the target population. A sample size of 100 to 200 participants was considered adequate for factor analysis (Comrey, 1978; Kline, 1986). A total of 350 participants including 153 males and 197 females completed the survey package in this study therefore the sample size was adequate to also conduct separate analysis according to sex (sample size after the removal of 2 participants with missing data). Second, the sample was required to represent the population of interest in order to determine the generalizability of the findings (Ferguson & Cox, 1993). The population of interest was young adults that were interested in physical activity but may or may not be physically active. According to the mean GLTEQ scores, the men and women were considered physically active. The GPSCS responses were assessed for homogeneity of variance-covariance among men and women to identify whether it would be more appropriate to examine a factor solution for the sample as a whole or according to sex.

Exploratory factor analysis techniques require variables to demonstrate univariate normality (Brown, 2006). Currently in EFA no ‘real’ guidelines exist to manage data of varying degrees of skewness and kurtosis (Ferguson & Cox, 1993). In larger samples, skewness does not make a substantive difference in the analysis (Tabachnick & Fidell, 2007).

The adequacy of the factor solution was assessed according to an inspection of the scree plot, a Monte Carlo parallel analysis, and the factor pattern of the factor coefficients. A scree test was conducted to determine the number of factors to retain in the factor solution (Cattell, 1966). Scree tests plot the eigenvalues for each factor against the number of factors (Ferguson & Cox, 1993). During the inspection of the scree plot, the point at which a change in slope is found identifies the number of factors (Cattell, 1978).

Parallel analysis was the second method used to identify the number of factors. Parallel analysis, such as Monte Carlo parallel analysis, compares a randomly produced set of eigenvalues generated with the same sample size as the observed values with those produced by the observed scores (Ferguson & Cox, 1993). Exploratory factor analysis was performed using this random data set to produce eigenvalues which were averaged for each factor. The average random eigenvalue was compared to the observed results. The number of eigenvalues from the observed data set that were greater than the corresponding values from the randomly-generate eigenvalues indicated the number of factors.

The third method of interpreting the number of factors was an examination of the factor patterns for simple structure (Thurstone, 1947). Simple structure is observed when items load highly on only one factor (Thurstone, 1947). Factor loadings > .30 serve as the lower boundary of a ‘meaningful’ factor loading (Gorsuch, 1983).

Subsequently, the factorability of the matrix was examined using the Kaiser-Meyer-Olkin (KMO) and Bartlett test of sphericity (BS). A KMO value of >.60 and a significant BS value supported the factorability of the matrix (Tabachnick & Fidell, 2007). The KMO looked for associations between variables in the correlation matrix, while the BS examined the lack of relationship existing between the variables (Ferguson & Cox, 1993).

Results

Preliminary data analyses

The rate of missing data was extremely low in this sample (GPSCS missing data < .06%). Respondents were excluded from the analysis if GPSCS responses were missing. Two male participants were subsequently excluded from the data set due to missing data in the GPSCS. To determine if it was suitable to combine the male and female data sets into a single data set for factor analysis purposes, the homogeneity of the variance-covariance matrices for male and female responses was assessed using Box's *M* test. A significant Box's *M* test statistic was obtained with an $F(120{,}331660) = 2.07$, $p < .001$ which suggests heterogeneity of the data. Consequently, EFA's were conducted separately on the male and female GPSCS data sets.

Exploratory factor analyses. The data were subjected to further pre-screening to assess the appropriateness of EFA. The data from the male sample and the female sample were considered suitable for factor analysis as the Bartlett's Test of Sphericity were significant ($p < .001$) and the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) supported the factorability of the matrix (KMO_{male}=.88 and KMO_{female}= .89) (Tabachnick & Fidell, 2007 p. 614).

The scree test suggested the retention of a 2-factor or 4-factor solution for men and a 3-factor or 4-factor solution for women (see Figure 4.1 and Figure 4.2). The Monte Carlo PCA for parallel analysis indicated the retention of three factors for both samples (Lautenschlager, 1989; see Table 4.1). Given the results obtained by the scree test and parallel analyses, 3-factor solutions were retained for both the male and female data sets.[4]

Following the direct oblimin rotations, the 3-factor solutions showed various levels of inter-correlation between factors (see Table 4.4). The pattern matrices for both men and women supported a 3-factor solution with a muscle-

[4] 2-factor and 4-factor solutions for each data set were examined, but the 3-factor solution provided the best theoretical interpretability and simple structure.

bulk/muscle-tone factor, a body-fat/muscle-tone factor (negative coefficients for muscle-tone), and a muscle-strength factor. The pattern matrices identified numerous problematic items (Item 1, 2, 3, 5, 8, 11, 15, 17) that lacked simple structure (see Table 4.2). The factor solutions did not indicate good discrimination between the factors as the muscle-tone items generally cross-loaded onto the muscle-bulk or body-fat factors (see Table 4.2 for results). The poor discrimination of muscle-tone items resulted in the re-analysis of the data omitting these items.

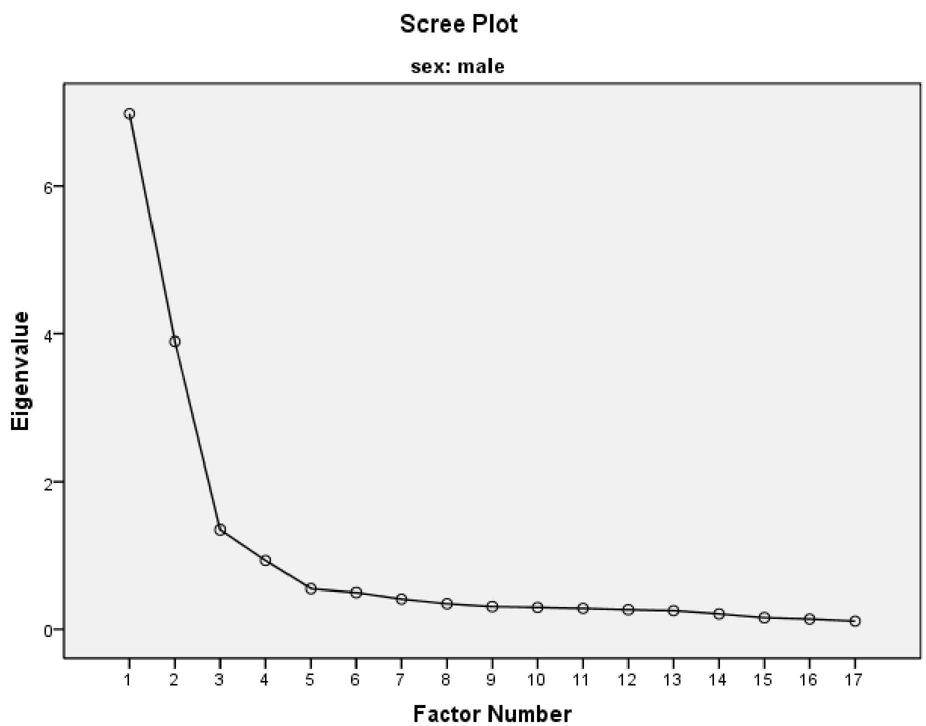

Figure 4.1. Scree plot of the eigenvalues for GPSCS for the sample of male university students.

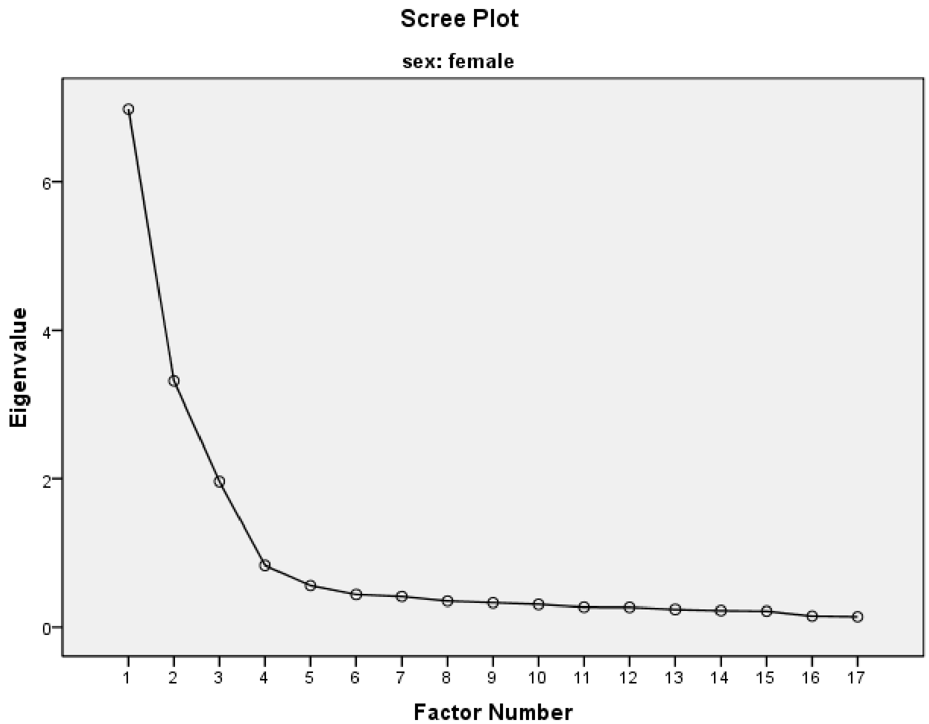

Figure 4.2. Scree plot of the eigenvalues for GPSCS for the sample of female university students.

Table 4.1

Eigenvalues from the Exploratory Factor Analysis (EFA) of GPSCS Data (17 items) and Corresponding Parallel Analysis

	Eigenvalues			
	Men		Women	
Factor	EFA	Parallel analysis	EFA	Parallel analysis
1	6.98	1.60	6.98	1.52
2	3.90	1.47	3.32	1.41
3	1.35	1.35	1.96	1.32
4	.93	1.28	.83	1.25

Table 4.2

Pattern Coefficients for Oblimin Principal Axis Solution of GPSCS Data (17-items) in Men and Women

		Factor					
		Men			Women		
Item #	Intended domain	1	2	3	1	2	3
17	MT	.80			.47	-.40	
14	MT	.77			.67		
11	MT	.67	-.40		.50	-.45	
6	MB	.66			.78		
5	MT	.65			.42	-.53	
15	MB	.65	.40		.82		
1	MT	.64			.39	-.63	
2	MB	.54		-.37	.70		
8	MB	.51	.46		.77		
10	BF		.88			.86	
4	BF		.86			.94	
16	BF		.83			.85	
13	BF		.81			.79	
12*	MS			-.88			.84
7*	MS			-.78			.90
9	MS			-.67			.71
3	MS	.31		-.61			.76

Note: Pattern coefficients > |.30| are presented in the matrix. MT= Muscle-tone; MB= Muscle-bulk; MS= Muscle-strength; BF= Body-fat. Content for each item number is presented in Appendix I.

Table 4.3

Inter-factor Correlations of the GPSCS (17 items) in Men and Women

	Factor		
Factor	1	2	3
1	-	-.09	-.58
2	-.17	-	0.02
3	.37	-.30	-

Note: Factor 1= Muscle-tone and Muscle-bulk, Factor 2 = Body-fat and Muscle-tone, Factor 3 = Muscle-strength. Values in the upper triangular matrix represent inter-factor correlations in men. Values in the lower triangular matrix represent inter-factor correlations in women.

Exploratory factor analysis of the muscle-bulk, muscle-strength and body-fat items of the GPSCS (following the removal of the muscle-tone items). After removing the four items that were originally intended to measure muscle tone, the inter-item correlation matrix for the 12 remaining items was re-analyzed using principal axes analysis with a direct oblimin transformation (delta = 0). The Bartlett's Test of Sphericity was significant ($p < .001$) and the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) supported the factorability of the matrices (KMO_{male}=.84 and KMO_{female}= .82) (Tabachnick & Fidell, 2007 p. 614).

The scree test (see Figure 4.3 and Figure 4.4) suggested a 3 factor solution for both samples. The Monte Carlo PCA for parallel analysis (Lautenschlager, 1989; see Table 4.4) suggested a 3 factor solution for the sample of women and a 2 factor solution for the sample of men. Given the results obtained by the scree tests and parallel analysis, 3-factor solutions were retained for both male and female data sets.[5]

The pattern matrices (see Table 4.4) supported the retention of 3-factor solutions for men and women. Among men, one muscle-bulk item (Item 2) had a cross-loading (pattern coefficient= .38) on the muscle-strength factor. Otherwise simple structure across all items for men and women were observed. Following the oblimin rotation, the three factors showed various levels of inter-correlation (Table 4.6 contains the inter-factor correlations). The three factors among both men and women were distinguishable from each other. Among men, the body-fat factor had low correlations with muscle-strength (r = -.17) and muscle-bulk (r = -.17). Muscle-bulk and muscle-strength had a strong correlation (r = -.55). These findings suggest that the body-fat factor is unrelated to both the muscle-bulk and muscle-strengths factors, while muscle-bulk and muscle-strength are more closely related. Among women, all three factors had lower inter-factor correlations ($r_{\textit{muscle-strength and body-fat}}$= -.23, $r_{\text{body-fat and muscle-bulk}}$= -.02, $r_{\textit{muscle-strength and muscle-bulk}}$ = -.33) which provides further evidence of distinguishable constructs.

[5] 2-factor and 3-factor solutions were examined for men. The 2-factor solution included muscle-bulk, muscle-tone and muscle-strength items loading onto the first factor; and the body-fat items, one muscle-bulk item and two muscle-tone items loading onto the second factor. Five cross-loaded items were observed in the 2-factor solution. A benefit of the 3-factor solution was the interpretability of the solution. Factor 1 included the muscle-bulk and muscle-tone items and one muscle-strength item. Factor 2 included the body-fat items, one muscle-tone item and two muscle-bulk items. Factor 3 included muscle-strength items and one-muscle-bulk item. A second benefit of the 3-factor solution was the similar overall factor pattern among the men and female samples.

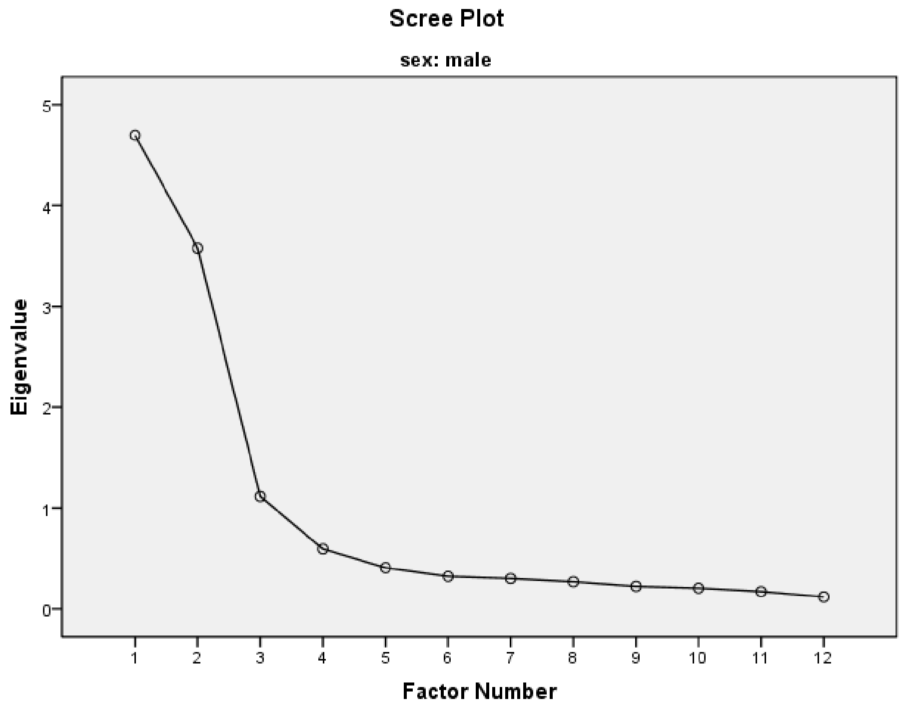

Figure 4.3. Scree plot of the eigenvalues for 12 GPSCS items (muscle-bulk, muscle-strength and body-fat items) for the sample of male university students.

Exploratory factor analysis of the muscle-tone items of the GPSC. The responses to the five muscle-tone items that had been removed from the GPSCS in the previous analysis were subjected to principal axis analysis with direct oblimin rotation. The scree test (see Figure 4-5 and Figure 4-6) and the Monte Carlo PCA for parallel analysis (Lautenschlager, 1989, see Table 4.7) determined that one factor should be retained for both male and female samples.

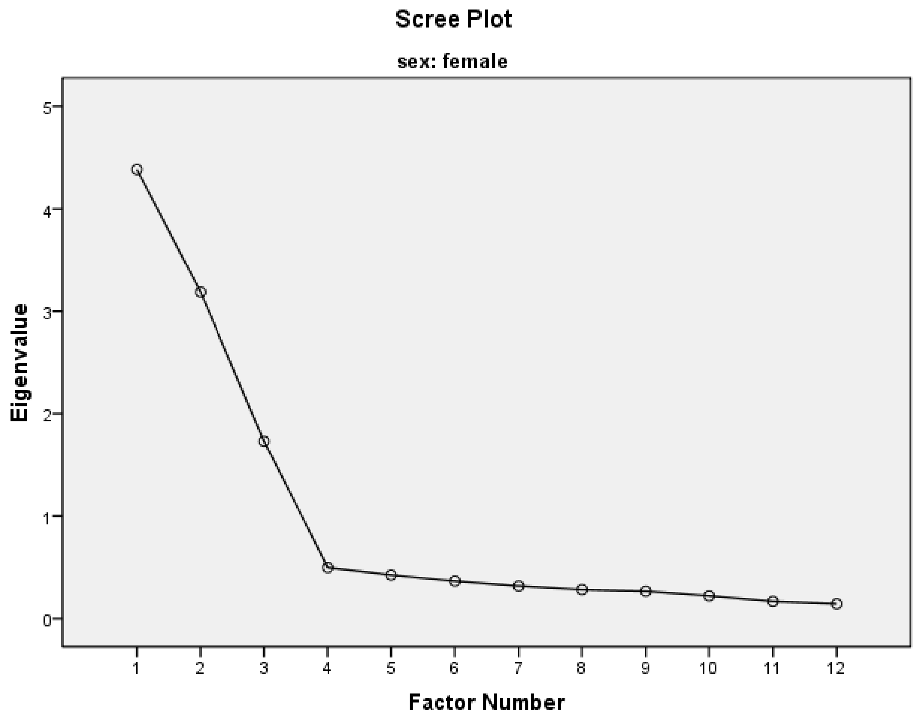

Figure 4.4. Scree plot of the Eigenvalues for 12 GPSCS items (muscle-bulk, muscle-strength and body-fat items) for the sample of female university students.

Table 4.4

Eigenvalues from the Exploratory Factor Analysis (EFA) of GPSCS Data (12 items) and Corresponding Parallel Analysis

	Eigenvalues			
	Men		Women	
Factor	EFA	Parallel analysis	EFA	Parallel analysis
1	4.70	1.49	4.39	1.42
2	3.58	1.35	3.19	1.31
3	1.12	1.25	1.73	1.22
4	.59	1.16	.50	1.15

Table 4.5

Pattern Coefficients for Oblimin Principal Axis Solution of GPSCS Data (12 items) in Men and Women.

		Factor					
		Men			Women		
Item #	Intended domain	1	2	3	1	2	3
12*	MS	.85			.82		
9	MS	.77			.72		
7*	MS	.74			.88		
3	MS	.69			.77		
10	BF		.91			.86	
16	BF		.90			.86	
13	BF		.86			.81	
4	BF		.86			.91	
6	MB			-.88			-.81
15	MB			-.82			-.81
8	MB			-.78			-.80
2	MB	.38		-.57			-.70

Note: Pattern coefficients > .30 were presented in the matrix. MB= Muscle-bulk; MS= Muscle-strength; BF= Body-fat. See Appendix I for a list of the item numbers and content.

Table 4.6

Inter-factor Correlations of the GPSCS (12 items) in Men and Women

	Factor		
Factor	1	2	3
1	-	-.17	-.55
2	-.23	-	-.17
3	-.33	-.02	-

Note: Factor 1= Muscle-strength, Factor 2 = Body-fat, Factor 3 = Muscle-bulk. Values in the upper triangular matrix represent inter-factor correlations in men. Values in the lower triangular matrix represent inter-factor correlations in women.

Bartlett's Tests of Sphericity were significant ($p < .001$) and the Kaiser-Meyer-Olkin measure of sampling adequacy supported the factorability of the matrices ($KMO_{male} = .88$ and $KMO_{female} = .85$) (Tabachnick & Fidell, 2007 p. 614). The loadings of items on the factor were high among men and women (factor loadings $_{men}$ = .75 to .83; factor loadings $_{women}$ = . 66 to .88) for the single factor (see Table 4.8). These findings suggest a single factor is present.

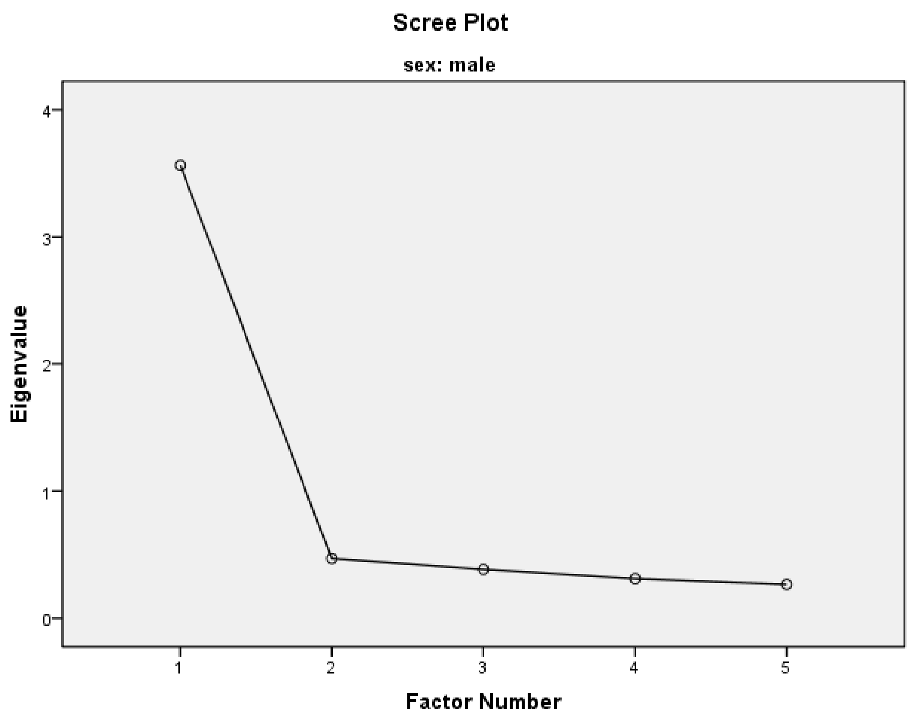

Figure 4.5. Scree plot of the eigenvalues for 5 GPSCS items (muscle-tone items) for the sample of male University students.

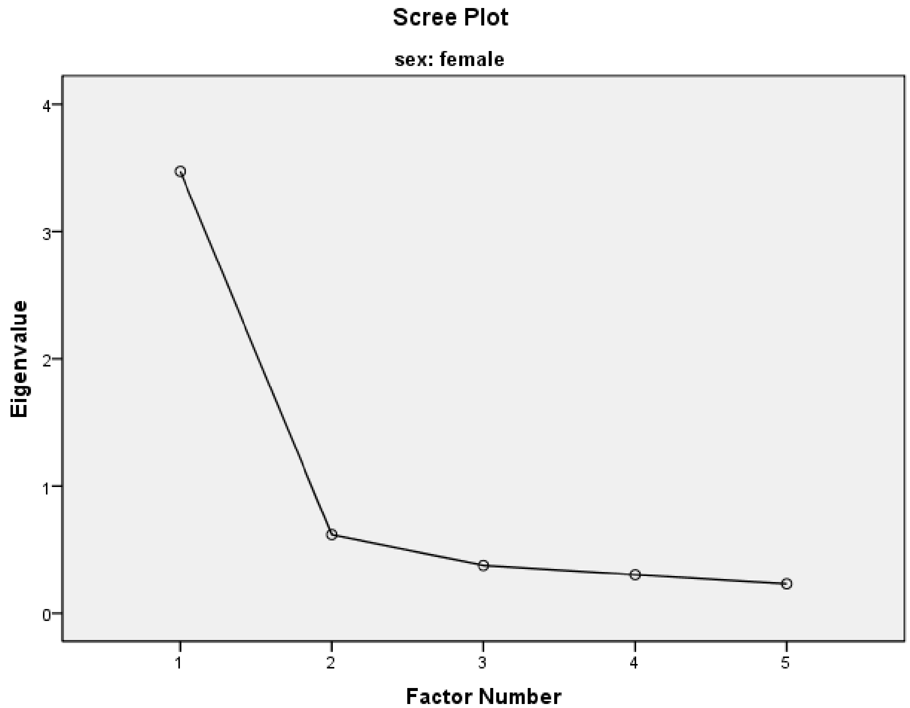

Figure 4.6. Scree plot of the eigenvalues for 5 GPSCS items (muscle-tone items) for the sample of female university students.

Table 4.7

Eigenvalues from the Exploratory Factor Analysis (EFA) of GPSCS Data (muscle-tone items, 5 items) and Corresponding Parallel Analysis

	Eigenvalues			
	Men		Women	
Factor	EFA	Parallel analysis	EFA	Parallel analysis
1	3.56	1.22	3.47	1.20
2	0.47	1.10	0.62	1.09

Table 4.8

Factor Loadings for Single-Factor Principal Axis Factoring Solution of Muscle-Tone Items Removed from GPSCS in Samples of Men and Women.

		Factor	
		Male	Female
Item #	Intended domain	1	1
11	MT	.83	.88
17	MT	.87	.80
5	MT	.78	.80
1	MT	.79	.78
14	MT	.75	.66

Note: MT= Muscle-tone. See Appendix I for a list of the item content according to item number.

Table 4.9

Global Physical Self-Concept Scale Subscale Means, Standard Deviations, and Internal Consistency Estimates Among Male and Female University Students.

	Men				Women	
Constructs	*M*	*SD*	α	*M*	*SD*	α
Muscle-tone	3.11	.80	.90	2.45	.80	.89
Muscle-bulk	2.58	.88	.89	1.81	.80	.87
Muscle-strength	4.11	.69	.86	3.51	.87	.88
Body-fat	1.77	.92	.89	2.59	1.14	.87

All subscales had acceptable levels of internal consistency (αs ≥ .86) (see Table 4.9 for means, standard deviations and internal consistency of GPSCS subscales). A one-way MANOVA was conducted to determine if gender differences existed across the four subscales of the GPSCS. A significant multivariate test statistic was obtained, Wilks' $\Lambda = .72$, $F(4, 345) = 63.00$, $p < .001$, $\eta^2 = .42$. Follow-up univariate *F*-test revealed significant gender differences on all four subscales: Muscle-strengh $F(1,348) = 47.69$, $p < .001$, $\eta^2 = .12$, Body-fat $F(1,348) = 52.71$, $p < .001$, $\eta^2 = .13$, Muscle-bulk $F(1,348) = 73.83$, $p < .001$, $\eta^2 = .18$, and Muscle-tone $F(1,348) = 59.03$, $p < .001$, $\eta^2 = .15$. Women scored significantly lower than the men in all subscales except body-fat on which they were significantly higher.

The GPSCS data (17 items) demonstrated a 3-factor solution with the muscle-tone items cross-loading on body-fat and muscle-tone. After removing the muscle-tone items, the GPSCS data for the muscle-bulk, body-fat and muscle-strength constructs were able to produce a clean 3-factor solution. In a separate analysis the GPSCS muscle-tone data produced a clean single factor solution.

Discussion

The purpose of this study was to develop structural validity evidence for the GPSCS (muscle-tone, muscle-bulk, muscle-strength and body-fat). According to initial EFA results, three factors emerged from the GPSCS data when all 17 items were included in the analyses for male and female data sets. The first factor contained muscle-bulk and muscle-tone items, the second factor contained muscle-tone and body-fat items, and the third factor consisted solely of muscle-strength items (see Table 4.2). Many of the muscle-tone items were identified as being problematic with respect to the factorial structure of the 17-item GPSCS given the strong cross-loadings that these items had on multiple factors. These results indicate that within these samples of male and female University students, muscle-tone has a strong association with body-fatness and muscle-bulk. Self-perception of muscle-tone was defined as the perceived degree to which muscles look lean, firm and defined. In order to consider oneself to have a high degree of muscle-tone one needs to have some muscle development, as well as have a low degree of fat on the body to see the muscle definition. For these reasons it seems logical that a muscle-tone factor is not distinguishable from muscle-bulk or body-fat within the context of the GPSCS.

Two separate EFA re-analyses of the GPSCS items were conducted with 12 items representing muscle-bulk, muscle-strength and body-fat. As indicated in Table 4.5, muscle-bulk, muscle-strength and body-fat items had meaningful loadings only on the factors/constructs they were intended to measure for both men and women. These simple-structure patterns provided evidence that the muscle-bulk, muscle-strength and body-fat items supported the retention of a 3-factor solution among men and women. Two separate EFA re-analysis of the GPSCS muscle-tone items were conducted with the 5 items among men and women. As indicated in Table 4.8, the muscle-tone items had strong and meaningful loadings on a single factor for both men and women. The factor analyses demonstrated that muscle-tone, muscle-bulk, muscle-strength and body-fat items have high factor loadings on their intended factors when the muscle-tone items are analyzed separately.

Previous studies have reported that the ideal male body consists of a mesomorphic build, with a defined upper body, muscular arms, pectorals and shoulders, with a slim waist, hips and butt (McCreary & Sasse, 2000; Morrison, Morrison, & Hopkins, 2003; Ridgeway & Tylka, 2005; Weinke, 1998). Whereas, an ideal female body consists of a thin and slender, as well as firm and toned body (Choi, 2000; Gruber, 2007; Markula, 1995). These sex differences in ideal physique may contribute to the structural validity of the GPSCS. The initial EFA with all GPSCS items found that among men the muscle-tone items factored onto the muscle-bulk factor while among women the muscle-tone items factored onto the muscle-bulk and body-fat (negative eigenvalue) factors. These differences may relate to sex differences in the conception of muscle-tone and differences in body ideals.

Qualitative research has clearly identified that women view having toned muscles differently than having bulky muscles, with muscle-bulk being an undesirable look even among female athletes (Choi, 2003; Markula, 1995; Mosewich, Vangool, Kowalski, & McHugh, 2009). In contrast, men desire bulky muscles such as large, broad shoulders and back, a thick neck, and a big chest and biceps; and a toned lower body which includes their butt, thighs and calves (Ridgeway & Tylka, 2005; Thompson & Cafri, 2007). For males they may possess muscle-bulk in their upper body while having a toned lower body. Among men, the general measure of muscle-bulk and muscle-tone may be problematic, as they may possess a bulky upper body and a toned lower body. Items that incorporate body site specificity may be required to distinguish between muscle-tone and muscle-bulk among men. The structural validity issues associated with muscle-tone and muscle-bulk in the GPSCS may be due to gender differences associated with ideal physiques and the generality of the instrument (not body site specific).

According to the MANOVA that was conducted to determine if sex differences existed across the GPSCS subscales, men and women had significantly different perceptions of muscle-tone, muscle-bulk, muscle-strength and body-fat (see Table 4.9). Men scored higher on perceptions of muscle-tone, muscle-bulk and muscle-strength and lower on body-fat than women. These findings are interpretable according to gender-specific body-image studies (see Choi, 2003; Markula, 1995; Mosewich, Vangool, Kowalski, & McHugh, 2009; Ridgeway & Tylka, 2005; Thompson & Cafri, 2007). As previously mentioned, men commonly focus on developing a large upper-body, a defined abdominal area and a toned lower body (Ridgeway & Tylka, 2005). However, women pursue the development of a thin, slender and firm body (Gruber, 2007; Markula, 1995). These sex-differences in body ideals and GPSCS scores indicate that separate analyses should be conducted among men and women in order to understand the roles of PSC of muscle-tone, muscle-bulk, muscle-strength and body-fat.

The combination of Studies 1, 2 and 3 establish initial validity evidence supporting the GPSCS as a measure of PSC of muscle-strength, muscle-bulk, muscle-tone and body-fat. Study 3 suggests that muscle-bulk, muscle-strength and body-fat subscales are three distinct factors according to factorial validity evidence. Muscle-tone items were not consistently supported. Content relevance and content representativeness evidence supported the muscle-tone items and subscale from studies 1 and 2. Factorial validity evidence was only established when muscle tone items were assessed independently from the GPSCS items. The mean subscale scores were significantly different between sexes and interpretable in a theoretically meaningful manner with men scoring higher on muscle-tone, muscle-bulk and muscle-strength and lower on body-fat than women.

Theoretically, these results support previous qualitative studies that have identified that men focus on the development of muscle-tone, muscle-bulk and muscle-strength (Ridgeway & Tylke, 2005) while women attempt to avoid developing too much muscle (Mosewich et al., 2009). This combination of validity evidence supports the GPSCS but also calls for further examination of muscle-tone.

One major concern of this study was the indistinguishability of the muscle-tone items within the context of the other GPSCS items according to the factor analyses in men and women. Muscle-tone items cross-loaded onto the muscle-bulk and body-fat factors in different patterns according to sex. The operational definitions of muscle-tone includes the appearance of lean, firm and defined muscles. This definition would require one to have some degree of muscle development which may explain the cross-loading of the muscle-tone items onto the muscle-bulk factor. Furthermore, the operational definition of muscle-tone identifies that muscles appear defined which would also require a low level of body-fat. This may explain why some muscle-tone items cross-loaded onto the body-fat factor. Further structural validity studies and criterion-related validity studies are necessary to understand and interpret the GPSCS responses among men and women. Further supporting evidence is required prior to the use of the GPSCS for the measurement of the PSC of muscle-tone, muscle-bulk, muscle-strength and body-fat.

Limitations

Limitations to this study include the homogeneity of the sample in terms of age, level of education and past experiences with physical activity. Further assessments in young adults from the general population beyond the University population would increase the generalizability of the findings. This group was a physically active sample, PSC may differ among active and inactive populations.

The process of construct validation is ongoing (Messick, 1995). The results from this study suggested that the GPSCS is a promising instrument for the assessment of muscle-tone, muscle-bulk, muscle-strength and body-fat. Further validation evidence is necessary prior to the use of the instrument. The potential for the GPSCS items to assess the constructs of muscle-tone, muscle-bulk, muscle-strength and body-fat have been supported but some limitations on the interpretation of the GPSCS have been introduced and will require further testing. Specifically, the interpretability of the muscle-tone construct among men and women needs further validity evidence to support the use within the GPSCS. The muscle-tone construct should be examined according to sex as men and women have different muscularity ideals and interpretations (Ridgeway & Tylka, 2005; Steinfeldt, Carter, Benton, & Steinfeldt, 2011). Further structural and criterion validity studies may assist in the understanding of the GPSCS constructs, the interpretability of the muscle-tone construct, gender differences in perceptions and interpretations of body physique, and limitations associated with the GPSCS.

References

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York, NY: Guilford Press.

Cattell, R. (1966). *Handbook of multivariate experimental psychology*. Chicago, IL: Rand McNally.

Cattell, R. (1978). *The scientific use of factor analysis in the behavioral and life sciences.* New York, NY: Plenum Press.

Choi, P. Y. L. (2003). Muscle matters: Maintaining visible differences between women and men. *Sexualities, Evolution and Gender, 5*, 71-81. doi:10.10 80/14616660310001632554

Comrey, A. (1978). Common methodological problems in factor analytical studies. *Journal of Consulting and Clinical Psychology, 46*, 648-659. doi:10.1037/0022-006X.46.4.648

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis.* Hillsdale, NJ: Erlbaum.

Dixon, B. J., Grimshaw, G. M., Linklater, W. L., & Dixson, A. F. (2011). Eye-tracking of men's preferences for waist-to-hip ratio and breast size of women. *Archives of Sexual Behavior*, *40*, 43-50. doi:10.1007/s10 508-009-9523-5

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but …: A meta-analysis of research on the physical attractiveness stereotype. *Psychological Bulletin, 110*, 109-128. doi:10. 1037/0033-2909.110.1.109

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin*, *111*, 304-341. doi:10.1037/0033-2909.111.2.304

Ferguson, E., & Cox, T. (1993). Exploratory factor analysis: A user's guide. *International Journal of Selection and Assessment*, *1*, 84-94. doi:10.1111/j.1468-2389.1993.tb00092.x

Fox, K. R. (2000). Self-esteem, self-presentations and exercise. *International Journal of Sport Psychology, 31*, 228-240.

García Bengoechea, E., Spence, J. C., & McGannon, K. (2005). Gender differences in perceived environmental correlates of physical activity. *International Journal of Behavioral Nutrition and Physical Activity, 2,* 12. doi:10.1186/1479-5868-2-12

Godin, G., & Shephard, R. J. (1997). Godin leisure time exercise questionnaire. *Medicine and Science in Sport and Exercise, 29*(6), S36-S38. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Gruber, A. J. (2007). A more muscular female body ideal. In J.K. Thompson & G. Cafri (Eds.), *The muscular ideal: Psychological, social, and medical perspectives* (pp. 217-234). doi:10.1037/11581-011

Jackson, L. A., Hunter, J. E., & Hodge, C. N. (1995). Physical attractiveness and intellectual competence: A meta-analytic review. *Social Psychology Quarterly, 58*(2), 108-122.

Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design.* New York, NY: Methuen and Co.

Lautenschlager, G. J. (1989). A comparison of alternatives to conducting Monte Carlo analyses for determining parallel analysis criteria. *Multivariate Behavioral Research, 24*, 365-395. doi:10.1207/s15327906mbr2403_6

Markula, P. (1995). Firm but shapely, fit but sexy, strong but thin: The postmodern aerobicized female body. *Sociology of Sport Journal, 12*(4), 424-453. Retrieved from http://journals.humankinetics.com/ssj

McCreary, D. R., & Sasse, D. K. (2000). An exploration of the drive for muscularity in adolescent boys and girls. *Journal of American College Health, 48*, 297-304. doi:10.1080/07448480009596271

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performance as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749. doi:10.1037/00 03-066X.50.9.741

Mosewich, A. D., Vangool, A. B., Kowalski, K. C., & McHugh, T. F. (2009). Exploring women track and field athletes' meanings of muscularity. *Journal of Applied Sport Psychology, 21*, 99-115. doi:10.1080/104 13200802575742

Morrison, T. G., Morrison, M. A., & Hopkins, C. (2003). Striving for bodily perfection? An expectation of the drive for muscularity in Canadian men. *Psychology of Men and Muscularity, 4*, 111-120. doi:10.1037/1524-92 20. 4.2.111

Nezlek, J. B. (1999). Body image and day-to-day social interaction. *Journal of Personality, 67*, 793-817. doi:10.1111/1467-6494.00073

Parks, F. R., & Kennedy, J. H. (2007). The impact of race, physical attractiveness, and gender on education majors' and teachers' perceptions of student competence. *Journal of Black Studies, 37*, 936-943. doi:10.1177/002193 4705285955

Peretti, P. O., & Abplanalp, R. R. (2004). Chemistry in the college dating process: Structure and function. *Social Behavior and Personality, 32*(2), 147-154. Retrieved from http://www.sbp-journal.com/index.php/sbp/index

Ridgeway, R. T., & Tylka, T. L. (2005). College men's perceptions of ideal body composition and shape. *Psychology of Men and Masculinity, 6*, 209-220. doi:10.1037/1524-9220.6.3.209

Steinfeldt, J. A., Carter, H., Benton, E., & Steinfeldt, M.C. (2011). Muscularity beliefs of female college student- athletes. *Sex Roles, 64*, 543-554. doi:1 0.1007/s11199-011-9935-2

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Toronto, ON: Pearson.

Thompson, J. K., Heinberg, L. J., Altabe, M., & Tantleff-Dunn, S. (2002). *Exacting beauty: Theory, assessment and treatment of body image disturbance.* Washington, DC: American Psychological Association.

Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago, IL: University of Chicago Press.

Weinke, C. (1998). Negotiating the male body: Men, masculinity, and cultural ideals. *Journal of Men's Studies, 6*(3), 255-282. Retrieved from http://www.mensstudies.info/wordpress/?page_id=46

Study 4

Validity and Reliability of the General Physical Self-Concept Scale in Female University Students

A critical aspect of psychometric testing is the development of multiple lines of evidence supporting the inferences made by test scores (Messick, 1989, 1995b). Messick (1995a, 1995b) recommends a unified concept of construct validity which addresses and integrates considerations of the content, criteria and the consequence. Sources of evidence in construct validity may include almost any kind of information about a test that is able to contribute to the understanding of the score meaning (Messick, 1995a). The combination of empirical evaluation and theoretical rationale underlying score interpretation provides a strong degree of evidence (Messick, 1995a). More specifically, evidence of construct validity may be developed by assessing: the expected performance of the construct; the differences between groups or settings; the findings in response to treatment or manipulations; the boundaries of the construct (Messick, 1989); the fidelity of the scoring structure according to the structure of the construct domain (Loevinger, 1957); the comparison with external factors (Messick, 1995a). The integration of score inferences that provide convergent evidence supporting the construct and discriminant evidence discounting plausible rival inferences contribute to the construct validity of psychological assessment (1995b).

The purpose of Study 4 was to examine psychometric properties of the GPSCS from an integrated construct validity approach among women. To address this purpose, internal structural validity, external validity (criterion-related validity) and reliability assessments of the GPSCS subscales as measures of physical self-concept were conducted. Internal structural validity evidence for the GPSCS was assessed using exploratory factor analysis (EFA). Internal consistency reliability (coefficient α; Cronbach, 1951) assessed the consistency of scores across items within the subscale (Aiken, 1985, p.88). The reliability of stability of the test scores within the same sample was examined over a one-week time period to assess test-retest reliability. External criterion-related validity of the GPSCS was assessed with respect to leisure-time physical activity participation (Godin & Shephard, 1997), body mass index (BMI), exercise identity (Anderson & Cychocsz, 1994), drive for muscularity (McCreary & Sasse, 2000) and drive for thinness (Garner & Olmstead, 1984).

The study sample comprised solely women for a number of reasons. Previous literature has questioned the similarity or difference in the conceptuality of muscle-tone among men and women (Smolak & Muren, 2008). Smolak and Muren questioned whether the drive for muscle-tone (drive for leanness) was a component of thinness, muscularity or a separate aspect of body image. Secondly, Study 3 in this dissertation found that the muscle-tone items loaded onto different factors among men and women. Among men, the muscle-tone items primarily loaded onto the muscle-bulk factor, while in the female sample the muscle-tone items had a positive loading on the muscle-bulk factor and a negative loading on the body-fat factor. These findings identified different conceptualizations of the muscle-tone items among men and women which promoted separate assessments according to gender. The third issue is a practical concern around sampling in the University classroom environment. The classes are predominantly female therefore collecting an adequate number of males for the proposed analytical techniques was not possible in the time frame for the dissertation.

Structural validity of the GPSCS was assessed in Study 3 using EFA. Among women, a 3-factor solution was observed when all 17-items were included in the EFA, with the perceptions of muscle-tone items cross-loading onto the muscle-bulk and body-fat (negative factor loading). After re-analysis of the muscle-bulk, muscle-strength and body-fat items separately from the muscle-tone items, the pattern matrix revealed a highly interpretable 3-factor solution with excellent simple structure across all items; and a strong single factor was noted among the muscle-tone items when assessed by themselves. In the current study, EFA will be used as an analytical tool to empirically assess the internal structural validity by examining the latent dimensionality of the 17 items (DeVellis, 1991). According to Gorsuch (1997), EFA may be used for confirmatory purposes whereby replication of the same factor solution across independent samples provides factorial validity evidence for the instrument.

The reliability of an instrument refers to the dependability, consistency, and the repeatability of the test scores for a particular population (Cheung & Yip, 2005). Internal consistency of the scale and stability of the scores across time are two different types of reliability. Internal consistency refers to the level of consistency among test items for the same construct (Trochim, 2006). Coefficient alpha assesses internal reliability of items according to a construct of interest which takes into consideration the variance attributed to the difference between items and participants (Cortina, 1993*)*. Test-retest reliability assesses the stability or consistency of scores on a measure over time (Trochim, 2006). To assess the stability test scores a sample is tested on two occasions and the correlation between the scores on each occasion is assessed and expressed as the reliability estimate. Developing reliability evidence is a critical aspect of instrument development (DeVellis, 2012).

External criterion-related validity infers support for the appropriateness of a test item or scale relative to an external criterion score or “gold standard” by appraisal of the degree to which an empirical relationship exists or not in a meaningful or theoretically consistent manner (DeVellis, 1991; Messick, 1995b). The degree of strength of the empirical relationship between the test scale and the external criterion is usually assessed in terms of correlations or regressions to infer criterion-related validity (Messick, 1989). The relationship between a scale and an external criterion can be predicted as having a positive or negative relationship. Convergent validity evidence is established when strong positive correlations are observed between two instruments measuring similar constructs (Campbell & Fiske, 1959). Discriminant validity evidence is obtained when low or zero correlations are obtained between scales theorized to measure dissimilar constructs.

In this study, external criterion-related validity evidence was gathered according to external criterion scores for body mass index (BMI), leisure-time physical activity participation (Godin & Shephard, 1997), exercise identity

(Anderson & Cychocsz, 1994), drive for muscularity (McCreary & Sasse, 2000) and drive for thinness (Garner & Olmstead, 1984). BMI is an index reflecting an individual's body weight relative to his/her height. A higher BMI indicates that people are heavier for their height, while a lower BMI indicates that people are lighter for their height. BMI was hypothesed to be positively correlated with the body-fat subscale[6] and muscle-bulk subscale, as the greater the amount of muscle or fat the heavier the individual will be. Physical activity and exercise levels were hypothesized to be positively correlated with perceptions of physical strength, muscle-tone and muscle-bulk, while being negatively correlated with body-fatness. More specifically, reports of participating in strength training activities were expected to be positively correlated with perceptions of physical strength, muscle-tone, and muscle-bulk. Muscle-strength, muscle-tone, and muscle-bulk were likely outcomes of physical activity participation, especially with strength training and cardiovascular training; therefore these subscales were hypothesized to be positively correlated. Reports of participating in cardiovascular activity are expected to be negatively correlated with perceptions of body-fat.

Other criterion variables being assessed with the GPSCS in this study included measures of exercise identity (Anderson & Cychosz, 1994), drive for muscularity (McCreary & Sasse, 2000) and the drive for thinness (Garner & Olmsted, 1984). Exercise identity is defined as the extent to which exercise is viewed as an integral part of the self-concept (Anderson & Cychocsz, 1994), and drive for muscularity assesses motivation to become more muscular (McCreary &

[6] Note: A high body-fat score represents a perception of "having too much body-fat," and a low body-fat score represents the perception of "not being fat."

Sasse, 2000). These two constructs, exercise identity and drive for muscularity, were hypothesized to be positively correlated with muscle-tone, bulk and strength. Self-perceptions of one's body (i.e., strength, muscle-tone, etc.) were hypothesized to be positively associated with one's exercise identity, as exercisers are expected to have greater strength and muscularity than non-exercisers. Drive for thinness was the third scale used as a criterion variable and is defined as peoples' motivation to lose weight (Garner & Olmsted, 1984). A negative correlation between drive for thinness and body-fat was expected, as well as a positive correlation between drive for thinness and muscle-tone. Self-perception of having a high degree of body-fat was expected to be related to an increased level of motivation to lose weight and be thinner. Although, high self-perceptions of muscle-tone would generally be associated with increased satisfaction with weight status and therefore not as strongly positively related to motivation to lose weight, high perceptions of muscle-tone might be positively associated with high physical activity levels. The convergent and discriminant correlation patterns of these three scales in relation to the GPSCS were included to develop external criterion related validity evidence to support the use of the GPSCS as a measure of PSC of muscle-tone, muscle-bulk, muscle-strength and body-fat.

The interpretation of the findings of the internal structural validity, external criterion related validity and reliability assessments can provide support that the GPSCS was serving the intended function. Developing reliability and validity evidence for a general measure of physical self-concept that includes muscularity will support the use of this instrument in future studies to further understand the nature of physical self-perceptions, and their relationship to exercise motivation and exercise participation.

Methods

Participant Characteristics

The results from Study 3 found different factor structures according to sex. For this study, we focused on the development of the GPSCS among women. Participants included 210 undergraduate student volunteers from two Canadian Universities. This was an independent sample from Study 3. The students were from various degree programs with education (*n* = 64), physical education and recreation (*n* = 63), and arts (*n* = 32) being the three most prevalent faculties represented among the participants. A total of 186 female students completed the survey package at time-1 and 174 female students participated at time-2. A total of 150 female participants completed both the time-1and time-2 survey package packages, missing participants only completed the time-1 survey packages or the time-2 survey package.

The students had a mean age of 20.4 years (*SD* = 3.84). The most predominant cultural ancestries of the participants were as follows: British Isles origins *n* = 36, 17.1%; North American origins *n* = 36, 17.1%; Western European origins *n* = 28, 13.3%; Eastern European origins *n* = 28, 13.3%; no other cultural ancestry represented 10% or greater of the valid percent. Body mass index (BMI) was calculated from self-reported height and weight (mean = 23.62 kg/m^2, *SD* = 7.11). On average the participants reported doing strength training 1.75 days per week (*SD* = 1.48), cardiovascular activities 3.03 days per week (*SD* = 1.57) and flexibility activities 2.43 days per week (*SD* = 1.98).

Instruments

Demographics. The demographic questions asked participants to indicate their sex, age, height, weight, faculty of study, and ethnic or cultural group of ancestry. Self-report height and weight were used to calculate body mass index (BMI). Cultural ancestry was categorized according to Statistics Canada (2008).

Type of physical activity. The types of physical activity items assessed the frequency of strength training, cardiovascular training and flexibility training. Participants were to consider the last 7-days and report the number of days (frequency) they participated in strength training, cardiovascular activities and flexibility activities.

Godin leisure-time exercise questionnaire (GLTEQ). The GLTEQ assessed self-reported leisure-time physical activity (Godin & Shephard, 1997). Participants considered an average 7-day period and recalled the number of times they participated in 15 minutes of strenuous, moderate or mild exercise during their free-time. Strenuous exercise was described as exercise that results in the heart beating rapidly while doing activities such as running, hockey, soccer, squash, etc. Moderate exercise was described as exercise that was not exhausting and included exercise such as walking, tennis, easy swimming, etc. Mild exercise was considered exercise that requires a minimal effort, such as yoga, bowling, golf, etc. The frequency of strenuous, moderate and mild exercise were multiplied by 9, 5 and 3 METS respectively and then summed.

General physical self-concept scale (GPSCS). The GPSCS contains the 17-items proposed to comprise four subscales. The four subscales measure muscle-tone (5-items), muscle-bulk (4-items), muscle-strength (4-items), and body-fat (4-items), as described in the previous study.

Exercise identity scale. *The Exercise Identity Scale (EIS)* (Anderson & Cychocsz, 1994) measured the extent to which a person views exercise as an integral part of his/her self-concept (Anderson & Cychocsz, 1994). The EIS includes nine items rated on 7-point Likert scales ranging from 1 (strongly disagree) to 7 (strongly agree). Anderson and Cychocsz (1994) developed the instrument to express the EIS score as the sum of the nine items. Factor analysis has confirmed the unidimensional nature of the instrument and Cronbach's alpha coefficients ranged between .94 to .95 in three unique samples (Anderson & Cychocsz, 1994; Anderson, Cychosz, & Franke, 2001). Wilson and Muon (2008) found that the exercise identity scale may be expressed as a 2-factor model which includes a 'role-identity' factor (Items 1,2 and 6) and 'exercise beliefs' factor (Items 3,4,5,7,8, and 9). In Wilson and Muon (2008), the χ^2 test was significant for the 9-item, 1-factor model (χ^2 =122.90, df = 27, p < .01), RMSEA = .12; SRMR = .06; CFI = .91 and the 2-factor model (χ^2 =69.53, df = 26, p < .01), RMSEA = .08; SRMR = .04; CFI = .96. These results demonstrated a better fit for the two-factor model. Both the single scale and two subscale models were used in this study.

Drive for muscularity scale (DMS). The DMS is a 15 item instrument that assessed the extent to which people desire a muscular body (McCreary &

Sasse, 2000). The DMS includes two subscales, drive for muscularity *attitudes* (7 items) and drive for muscularity *behaviours* (8 items). “I lift weights to build up muscle” is a sample item for the drive for muscularity behaviours subscale and “I wish I were more muscular” is a sample item for the drive for muscularity attitudes subscale (McCreary & Sasse, 2000). Items were assessed on 6-point Likert scales ranging from 1 (very much like me) to 6 (not at all like me). A higher score on this scale suggests a greater drive for muscularity. The drive for muscularity behaviours subscale included one item that may not be appropriate for all samples (“I think about taking steroids”) and can be removed at the discretion of the researcher (McCreary, 2007). This item was removed from the analysis for this study as it was deemed to be unsuitable for a sample of female university students. The internal consistency reliability has been good (α >.80) for men and women (Cafri & Thompson, 2004; Davis, Karvinen, & McCreary, 2005; McCreary, Saucier, & Coutenay, 2005). The test-retest reliability of the instrument has not been reported for women (Cafri & Thompson, 2004).

Drive for thinness. Drive for thinness was assessed using the ‘drive for thinness’ subscale of the Eating Disorder Inventory (EDI) (Garner & Olmstead, 1984). The *drive for thinness* subscale included seven items (e.g., “I think about dieting”). Each item was assessed on a 6-point Likert scale ranging from 1 (always) to 6 (never). The internal consistency has been found to be high (α =.88) and test-retest reliability was good (r=.88) (Zabinski, Calfas, Gehrman et al., 2001).

Intention to exercise. A declarative statement concerning intentions to participate in strength training, cardiovascular exercise and flexibility exercise over the next week were posed followed by a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree).

Procedures

Following ethical approval from the institutional research ethics boards[7], students were recruited to volunteer in the study. Participants were informed both orally and in writing that participation was voluntary and they were free to withdraw without consequence at any time. The instrument was distributed in a classroom setting. The first survey package included all the instruments listed in the instrument section. One week later the participants completed a second survey package in the classroom setting which only included the GLTEQ, GPSCS, the type of physical activity questions, and intentions to exercise.

The data treatment was identical to Study 3. The adequacy of the factor solution was assessed according to an inspection of the scree plot, Monte Carlo parallel analysis and the factor pattern coefficients. Subsequently, the factorability of the matrix was examined using the Kaiser-Meyer-Olkin (KMO) and Bartlett test of sphericity (BS).

Internal consistency reliability. To assess internal consistency of the subscales in the GPSCS coefficient α (Cronbach, 1951) was examined.

***Test-retest reliability*.** To assess the consistency or repeatability of the GPSCS, test-retest reliability was assessed (Pedhazur & Schmelkin, 1991). The

[7] This study was conducted at two Universities in the same Western Canadian city. Ethical approval from each institutional ethics board was obtained.

data from the GPSCS was assessed at two time points; a time interval of one week between the first and second assessments was scheduled. As recommended by Thomas, Nelson, and Silverman (2005), test-retest reliability was assessed by computing intra-class correlation coefficients (ICC) for each of the four subscales. The resulting ICC was conceptualized as a coefficient of test score stability over the two testing sessions (Thomas et al., 2005). Large positive statistically significant ICCs (ICC >.70) are indicative of good test score stability (Nunnally, 1978).

Results

Physical Activity Participation and Intentions

Leisure-time physical activity was assessed at time-1 and -2. At time-1, a mean of 80.8 METS (SD = 80.9) of leisure time physical activity was calculated and a mean of 76.6 (SD = 76.4) was calculated at time two. No significant difference was observed in physical activity participation between the two time points; $F(1,142) = .57$, $p = .45$; partial $\text{Eta}^2 = .004$.

Exploratory Factor Analysis of the GPSCS

The data were subjected to further pre-screening to assess the appropriateness of EFA. The data from the time one sample were considered suitable for factor analysis as the Bartlett's Test of Sphericity was significant ($p < .001$) and the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) supported the factorability of the matrix ($KMO_{\text{time one}} = .89$) (Tabachnick & Fidell, 2007, p. 614).

The scree test suggested the retention of a 3-factor solution (see Figure 5.1). The Monte Carlo PCA for parallel analysis also supported the retention of three factors (Lautenschlager, 1989; see Table 5.1).

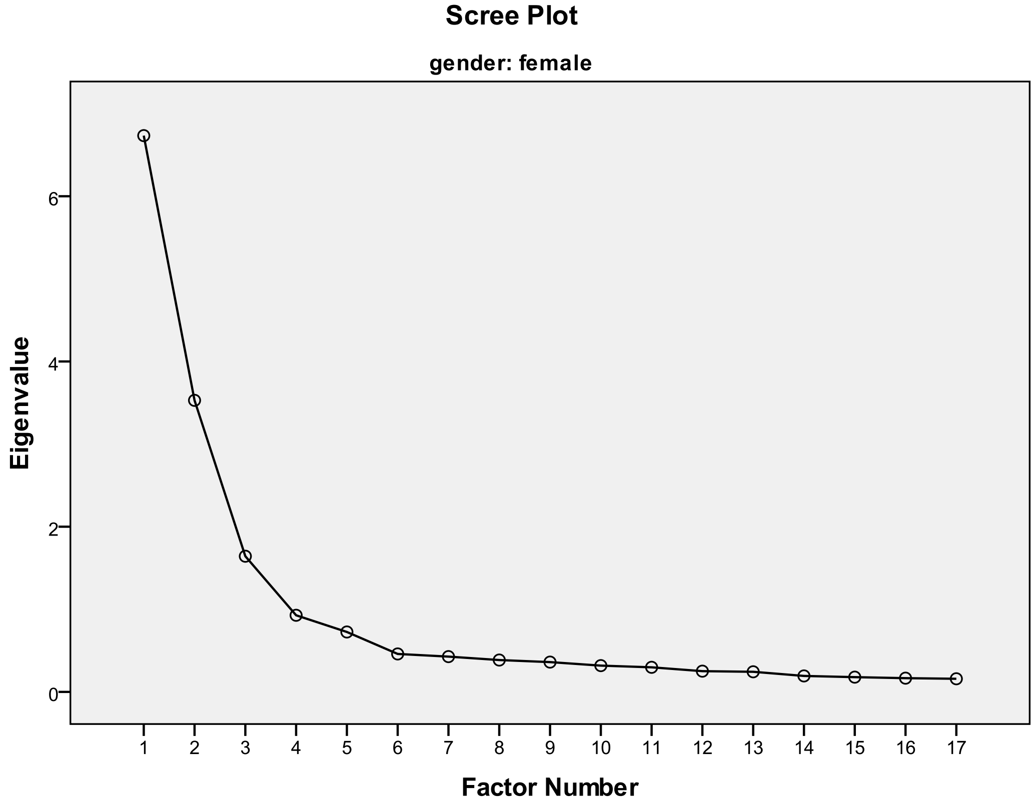

Figure 5.1. Scree plot of the eigenvalues for GPSCS for the sample of female university students.

The pattern matrix demonstrated a three factor solution with a muscle-bulk/muscle-tone factor (factor 1), a body-fat/muscle-tone factor (negative coefficients for muscle-tone; factor 2), and a muscle-strength factor (factor 3). The pattern matrix identified five problematic items (Items 1, 5, 9, 11 and 17) that lacked simple structure (see Table 5.2). The factor solution did not indicate good discrimination between the factors as the muscle-tone items generally loaded onto the muscle-bulk or body-fat factors (see Table 5.2). Poor discrimination of the muscle-tone items was observed which was consistent with the Study 3 findings. The data in Study 4 were re-analysed with the muscle-tone items removed.

Table 5.1

Eigenvalues from the Exploratory Factor Analysis (EFA) of GPSCS Time-1 Data (17-items and 12-items) and Corresponding Parallel Analysis

	Eigenvalue			
	17-item GPSCS		12-item GPSCS	
Factor	EFA	Parallel analysis	EFA	Parallel analysis
1	6.73	1.57	4.31	1.43
2	3.53	1.44	3.43	1.31
3	1.64	1.35	1.45	1.23
4	.93	1.27	.51	1.15

Note. The 17-item GPSCS comprises 5 muscle-tone items, 4 body-fat items, 4 muscle-strength items and 4 muscle-bulk items. The 12-item GPSCS includes the 4 body-fat items, 4 muscle-strength items and 4 muscle-bulk items.

Exploratory factor analysis of the muscle-bulk, muscle-strength and body-fat items of the GPSCS (removal of the muscle-tone items). After removing the five muscle-tone items, the data were re-analyzed using principal axes analysis with a direct oblimin transformation (delta = 0) with 12 items from the GPSCS (muscle-bulk, muscle-strength and body-fat items). The Bartlett's Test of Sphericity was significant ($p < .001$) and the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) supported the factorability of the matrices (KMO_{time-1}=.85 (Tabachnick & Fidell, 2007 p. 614).

Table 5.2

Pattern Coefficients for Oblimin Three-Factor Principal Axis Solution of GPSCS Time-1 Data in Women

		Factor		
Item #	Intended domain	1	2	3
15	MB	.77		
6	MB	.75		
2	MB	.72		
8	MB	.67		
14	MT	.53		
10	BF		.89	
4	BF		.89	
16	BF		.87	
13	BF		.75	
17	MT	.51	-.60	
1	MT	.44	-.57	
11	MT	.47	-.56	
5	MT	.35	-.51	
7*	MS			.89
12*	MS			.82
3	MS			.73
9	MS	.31		.61

Note: Pattern coefficients > |.30| are presented in the table. MT= Muscle-tone; MB= Muscle-bulk; MS= Muscle-strength; BF= Body-fat. See Appendix I for a list of the items. * = reverse scored items.

The scree test (see Figure 5.2) and the Monte Carlo PCA for parallel analysis (Table 5.1) (Lautenschlager, 1989; see Table 5.1) clearly suggested the retention of a 3-factor solution. Examination of the pattern matrix revealed excellent simple structure across all 12 items in the retained 3-factor solution (see Table 5.3). Table 5.4 contains the inter-factor correlations.

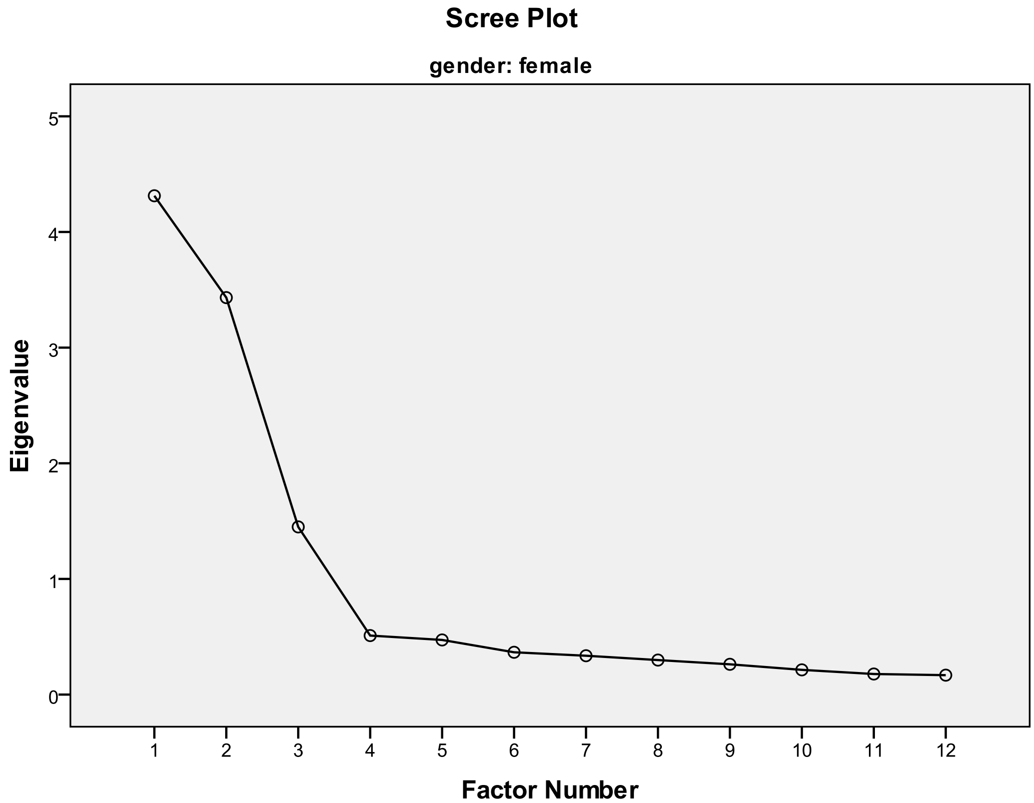

Figure 5.2. Scree plot of the eigenvalues for 12 GPSCS items (muscle-bulk, muscle-strength and body-fat items) for the sample of female university students.

Exploratory factor analysis of the muscle-tone items of the GPSCS (time one). The muscle-tone data were re-analyzed separately from the other GPSCS items using EFA. The data were subject to a principal axes analysis with direct oblimin rotation which was performed using PASW Statistics 19. Bartlett's Test of Sphericity was significant ($p < .001$) and the Kaiser-Meyer-Olkin measure of sampling adequacy supported the factorability of the matrices ($KMO_{\text{time one}}$ = .86) (Tabachnick & Fidell, 2007 p. 614). The scree test (see Figure 5.3) and the Monte Carlo PCA for parallel analysis (Lautenschlager, 1989, see Table 5.5) indicate that one factor should be retained.

Table 5.3

Pattern Coefficients for Oblimin Three-Factor Principal Axis Solution of GPSCS (muscle-bulk, muscle-strength and body-fat) Time-1 Data in Female University Students

		Factor		
Item #	Intended domain	1	2	3
7*	MS	.87		
12*	MS	.80		
3	MS	.74		
9	MS	.65		
16	BF		.91	
10	BF		.91	
4	BF		.89	
13	BF		.79	
6	MB			-.86
15	MB			-.77
8	MB			-.73
2	MB			-.71

Note. Pattern coefficients > |.30| are presented in the table. MB= Muscle-bulk; MS= Muscle-strength; BF= Body-fat. See Appendix I for a list of the items.

Table 5.4

Inter-factor Correlations of the GPSCS (12 items)

	Factor		
Factor	1	2	3
1	-	-	-
2	-.27	-	-
3	-.39	-.10	-

Note: Factor 1= Muscle-strength, Factor 2 = Body-fat, Factor 3 = Muscle-bulk.

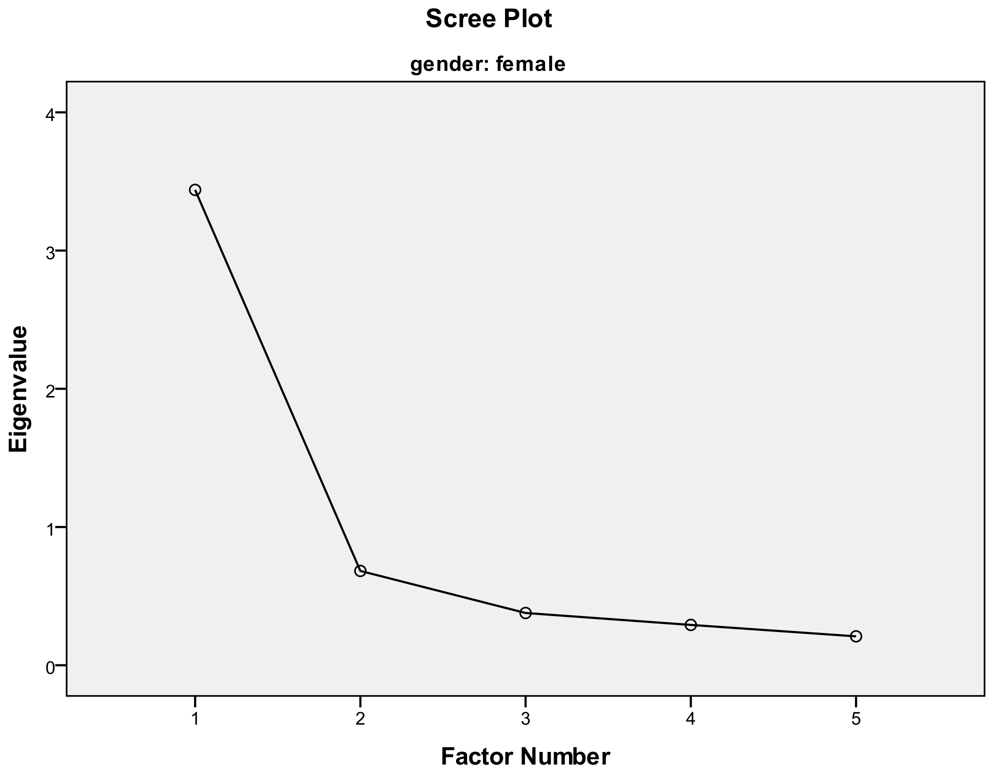

Figure 5.3. Scree plot of the eigenvalues for five GPSCS items (muscle-tone items) for the sample of female university students at time one.

Table 5.5

Eigenvalues from the Exploratory Factor Analysis (EFA) of GPSCS Time-1 Data (muscle-tone items) and Corresponding Parallel Analysis

	Eigenvalue	
Factor	EFA	Parallel analysis
1	3.44	1.21
2	0.68	1.09

Table 5.6

Factor Loadings for Single-Factor Principal Axis Factoring Solution of GPSCS Time-1 Data (muscle-tone items) in a Sample of Female University Students

		Factor
Item #	Intended domain	1
11	MT	.89
17	MT	.87
1	MT	.84
5	MT	.73
14	MT	.56

Note. MT= Muscle-tone. See Appendix I for a list of the items.

The EFA at time-1 replicated the findings from Study 3. The EFA results indicate that the muscle-bulk, muscle-strength and body-fat items demonstrate a clean simple structure when the muscle-tone items were omitted. When included in the exploratory factor analysis, the muscle-tone items cross-loaded on the body-fat and muscle-bulk factors. The muscle-tone items were re-analyzed separately and a single factor among female university students was observed. See Table 5.7 for the means and standard deviations for the GPSCS items at time-1 and time-2.

Reliability of the GPSCS

Internal consistency and test-retest reliability. Table 5.8 contains the subscale means, standard deviations, internal consistency coefficients (Cronbach's α; Cronbach, 1951), and the intra-class correlation coefficients. The internal consistency values across the four subscales at time-1 and time-2 were good (i.e., all αs $\geq .82$). The Time-1 17-item scale and the Time-2 17-item scale were used to calculate the intra-class correlation coefficients. All intra-class correlation coefficients were large (rs $\geq .75$; all ps $< .001$). It appears that the subscales of the

17-item GPSCS possessed excellent test-retest reliability over a 1-week period for the sample of undergraduate students. Overall, the GPSCS subscales demonstrated good internal consistency and stability over a 1-week time period in female University students.

Table 5.7

GPSCS Item Means, and Standard Deviations

	Time-1		Time-2	
Subscale and Item	*M*	(*SD*)	*M*	(*SD*)
Muscle-tone: The self-perception of the degree to which muscles appear lean, firm and defined.				
1. My muscles appear toned.	2.90	(.98)	2.98	(1.01)
5. My muscles appear firm.	2.87	(.86)	2.93	(.97)
11. My muscles appear well-defined.	2.55	(1.05)	2.64	(.99)
14. My muscles are ripped.	1.56	(.78)	1.80	(.85)
17. My muscles are well sculpted.	2.33	(.97)	2.51	(.99)

(continued)

	Time-1		Time-2	
Subscale and Item	*M*	(*SD*)	*M*	(*SD*)
Muscle-bulk: The self-perception of the degree to which muscles appear large and bulky in size.				
2. I appear to have large muscles.	2.19	(1.02)	2.27	(.94)
6. My muscles appear to be huge.	1.57	(.77)	1.72	(.77)
8. My muscles appear bulky.	1.67	(.82)	1.78	(.78)
15. My muscles appear big.	1.70	(.79)	1.90	(.82)
Physical strength: The self-perception of the presence or absence of muscle-strength and the ability to lift heavy objects.				
3. I am physically strong.	3.19	(1.03)	3.27	(.94)
7. I am physically weak.*	4.10	(.98)	4.10	(.98)
9. My muscles are strong.	3.17	(.93)	3.12	(.90)
12. My muscles are weak.*	3.85	(1.01)	3.90	(.94)
Body-fat: The self-perception of the degree of fat that appears on the body				
4. My body appears fat.	2.44	(1.14)	2.56	(1.12)
10. I appear chubby.	2.31	(1.24)	2.34	(1.23)
13. My body jiggles.	2.76	(1.13)	2.76	(1.07)
16. My body is flabby.	2.35	(1.17)	2.42	(1.09)

Note. * = reverse-scored item.

Table 5.8

Descriptive Statistics, Internal Consistencies (α), and Test-Retest Reliabilities (Intra-Class Correlation Coefficients: ICC) for GPSCS Subscales Among Female University Students

	Time One			Time Two			
Subscales	*M*	*SD*	α	*M*	*SD*	α	ICC
Muscle-bulk	1.78	(.72)	.85	1.92	(.69)	.85	.75**
Muscle-strength	3.58	(.81)	.87	3.53	(.75)	.82	.85**
Body-fat	2.47	(1.08)	.93	2.53	(1.03)	.92	.90**
Muscle-tone	2.66	(.85)	.89	2.76	(.89)	.91	.84**

Note. ICC = intra-class correlation coefficients for the GPSCS at time-one and time-two. ** = $p < .001$

One-way Repeated Measures ANOVA of GPSCS Subscales

One-way repeated measures ANOVA reveals a significant difference between muscle-tone, muscle-bulk, muscle-strength and body-fat subscales at time-1 (Wilks' $\Lambda = .16$, $F(3, 183) = 318.87$, $p < .001$, $\eta^2 = .84$). Post-hoc pairwise Bonferroni-corrected comparisons reveal that all GPSCS subscale pairings were significantly different ($p < .001$) except the body fat- muscle-tone pairing ($p = .69$).

Bivariate Correlations Between GPSCS Subscales

Bivariate correlations were calculated among the GPSCS subscales at time-1 and time-2 (Table 5.9). Among the GPSCS, muscle-tone has a significant positive relationship with muscle-bulk and muscle-strength, and an inverse relationship with body-fat at time-1 and time-2. The muscle-bulk subscale has a significant positive relationship with muscle-tone and muscle-strength, and no relationship with body-fat at time-1 and time-2. The muscle-strength subscale has a significant positive relationship with muscle-tone and muscle-bulk, and a significant inverse relationship with body-fat at time-1 and time-2. The body-fat subscale has a significant inverse relationship with muscle-tone and muscle-strength, and no relationship with muscle-bulk at time-1 and time-2.

Table 5.9

Bivariate Correlations Between GPSCS Subscales at Time-1 and Time-2

	Subscales			
Subscales	MT	MB	MS	BF
MT	-	.44**	.55**	-.54**
MB	.44**	-	.41**	.00
MS	.50**	.44**	-	-.35**
BF	-.54**	.06	-.23**	-

Note: ** $p \leq .01$ (2-tailed). MT = Muscle-Tone; MB = Muscle-Bulk; MS = Muscle-Strength; BF = Body-Fat. The lower-triangle of the correlation matrix includes the Time-1 subscales. The upper-triangle of the correlation matrix includes the time-2 subscales. $n_{Time\text{-}1} = 186$; $n_{Time\text{-}2} = 174$.

External Validity Evidence for the GPSCS

Criterion-related validity evidence. Criterion-related validity evidence was obtained for GPSCS subscales at time-1 when subscales related in theoretically meaningful ways with BMI, physical activity, exercise identity, drive for muscularity and drive for thinness. Means, standard deviations and alpha coefficients for all criterion variables are listed in Table 5.10.

Table 5.10

Descriptive Statistics and Bivariate Correlations Between BMI, GLTEQ, Type of Physical Activity, EIS, DMS and the GPSCS Subscales.

					Bivariate correlations			
Study Variables	*M*	*SD*	*n*	α	Muscle-tone	Muscle-bulk	Muscle-strength	Body-fat
BMI	23.62	(7.11)	185	-	-.19*	.10	.06	.42**
GLTEQ	80.77	(80.98)	180	-	.21**	.23**	.21**	-.07
Strength Training	1.75	(1.48)	185	-	.34**	.23**	.40**	-24**
Cardiovascular Activity	3.03	(1.57)	185	-	.32**	.17*	.33**	-.10
Flexibility Activity	2.43	(1.98)	185	-	.17*	.03	.25**	-.15*
EIS- Role Identity	4.38	(1.83)	186	.86	.50**	.36**	.46**	-.25**
EIS-Exercise Beliefs	4.74	(1.57)	186	.93	.40**	.27**	.36**	-.12
EIS- Sum of all EIS	41.58	(14.24)	186	.94	.46**	.32**	.42**	-.18*
DMS – Behaviour	4.90	(.75)	186	.76	-.32**	-.18*	-.33**	.13
DMS – Attitudes	4.27	(.84)	186	.81	.16*	.14	.19**	-.20**
EDI – Drive for Thinness	3.57	(1.11)	186	.83	.30**	.03	.20**	-.50**

Note. EIS = Exercise Identity Scale (Anderson & Cychocsz, 1994), DMS = Drive for Muscularity Scale (McCreary & Sasse, 2000), EDI = Eating Disorder Inventory (Garner & Olmstead, 1984), a lower EDI score suggests a greater endorsement of the drive for thinness; DMS= a high score suggest lower drive for muscularity and a lower DMS score indicates greater drive for muscularity; * $p < .05$; ** $p < .01$

An inspection of the bivariate correlations (Table 5.10) indicated, as hypothesized, that a higher BMI was associated with a higher endorsement of the body-fat scores. A lower BMI score was associated with stronger endorsement of self-perceptions of having a toned body. Higher self-reported levels of physical activity, according to the GLTEQ score, were associated with more positive perceptions of muscle-bulk, muscle-tone and physical strength. No relation between GLTEQ and body-fat was observed. Greater frequency of strength training and cardiovascular training were positively associated with self-perceptions of muscle-tone, muscle-bulk and physical strength. Greater frequencies of flexibility training were positively associated with self-reported perceptions of muscle-tone and physical strength. An inverse relationship was observed for strength and flexibility training according to self-perception of body-fat, which suggests lower physical activity levels (strength and flexibility) among participants with higher self-perceptions of having a fat body. Cardiovascular training was not associated with the degree of self-perception of body-fatness.

Further inspection of the bivariate correlations in Table 5.10 identified the relationship between physical self-concepts (GPSCS) and exercise identity, drive for muscularity and drive for thinness. A stronger endorsement of the exerciser identity (sum of EIS scores) was positively correlated with self-perceptions of muscle-tone, muscle-bulk and physical strength, and negatively correlated to perceptions of body-fat. The two factors of exercise identity within the EIS, EIS role-identity and EIS exercise beliefs, proposed by Wilson and Muon (2008) were assessed as well. Role identity was positively correlated to self-perceptions of muscle-tone, muscle-bulk and muscle-strength; and negatively correlated to self-perceptions of body-fat. Exercise beliefs were positively associated with self-perceptions of muscle-tone, muscle-bulk and strength; no significant relationship was found with self-perceptions of body-fat. The belief component of exercise identity includes the endorsement of having exercise goals, exercise being a part of the PSC, the need for exercise to feel good about oneself, and thinking about exercise often. It was expected that self-perceptions of having muscle-tone, muscle-bulk and muscle-strength would be related to exercise identity beliefs. Greater endorsement of behaviours associated with the drive for muscularity were reported by those with higher self-perceptions of muscle-tone, muscle-bulk and physical strength, these behaviours were not endorsed by participants with perceptions of higher body-fatness.[8] Greater endorsement of attitudes associated with the drive for muscularity reported by those with greater self-perceptions of body-fatness and lower perceptions of muscle-tone and physical strength. A greater drive for thinness, as assessed using the EDI subscale, was positively associated with positive perceptions of muscle-tone and muscle-strength. Drive for thinness and self-perceptions of body-fatness had a strong significant negative correlation. Collectively, the results of these bivariate correlation provide external criterion related validity evidence supporting the psychometric properties of the GPSCS subscales.

Discussion

The purpose of this study was to develop construct validity evidence for the GPSCS by examining the internal structural and external criterion-related validity in a sample of female university students. Exploratory factor analysis and internal reliability results supported:

---

[8] When using the drive for muscularity scale, a high score suggests lower DMS and a low score indicates greater DMS.

1. a 12-item, 3-factor instrument which measured the constructs of muscle-bulk, muscle-strength and body-fat (Part 1) and

2. a 5-item, single factor instrument which measured the construct of muscle-tone (Part 2) (See Appendix J for Part 1 and Part 2 of the GPSCS instrument).

Internal Structural Validity

The latent structure of the GPSCS items was assessed in this paper using EFA. The underlying processes in the GPSCS items were found to be complex. The four hypothesized constructs being assessed in the GPSCS included muscle-tone, muscle-bulk, muscle-strength and body-fat. When all 17 items of the GPSCS were included in the EFA, three, rather than four factors were found. The muscle-tone items loaded onto the same factor as the muscle-bulk items and the body-fat items (negative factor loadings) rather than expressing as an independent and distinguishable muscle-tone factor. With the removal of the muscle-tone items the other items factored out as hypothesized onto 3 distinct factors. A follow-up factor analysis with only the muscle-tone items demonstrated a strong, clean single factor solution. The Study 4 factor analyses replicate the results from Study 3 among women, which demonstrate that consistently over two samples of women, the muscle-tone items cannot be distinguished from the other GPSCS items when assessed in the same factor analysis. The inability of the muscle-tone items to demonstrate a clean factor among the GPSCS items may be due to the operational definition of muscle-tone construct including an aspect of having little body-fat and having some muscularity. The operational definition of muscle-tone is ‘the self-perception of the degree to which muscles appear lean, firm and defined’, which requires the body to have some muscle and have a low degree of body-fat to view the muscle definition.

Issues associated with ambiguity of the muscle-tone items have been an on-going concern. Study 2 asked fitness experts to recount how their clients described muscle-tone. The fitness experts in Study 2 identified that muscle-tone is a complex topic for clients. Some clients referred to muscle-tone as having very little body-fat which makes it possible to see muscle definition, while others referred to muscle-tone as the development of some muscle and a reduced amount of fat. Although these two conceptions of muscle-tone are very similar, there were some differences in training techniques and dietary intake to develop these two different types of muscle-tone. Fitness experts in Study 2 identified muscle-tone and muscle-bulk as distinct constructs, even though they used the terms “being ripped” and “sculpted” to define both constructs.

Smoklak and Murnen (2008) found that women were more invested in pursuing a drive for leanness than a drive for muscularity which support the ordinal trend found in the mean subscale score results (see Table 5.8). Women generally do not have large amounts of muscle-bulk unless they participate in exercise training that specifically focuses on increasing muscle mass. In this study, the mean subscale scores support this supposition in that women’s muscle-bulk scores were significantly lower than the other GPSCS mean subscale scores (see Table 5.8). The mean subscale scores provide information regarding women’s PSC which demonstrate the same ordinal trends at two time-points within Study 4 and in Study 3. The difference in mean muscle-tone and muscle-bulk scores provide further validity evidence supporting the distinction between these two constructs.

The bivariate correlations among subscales of the GPSCS demonstrated that muscle-tone has a different relationship with the GPSCS subscales than muscle-bulk or body-fat (see Table 5.9). The muscle-tone subscale has a significant inverse relationship with body-fat, while the muscle-bulk subscale has no relationship with body-fat. The distinct relationships between muscle-tone and body-fat, and muscle-bulk and body-fat is further evidence that muscle-tone and muscle-bulk are separate constructs.

A one-way repeated measure post-hoc pairwise Bonferroni-corrected comparisons revealed that absolute levels of the scores on the body-fat and muscle-tone subscales were not significantly different ($p = .69$). These results highlight the strong inverse relationship between body-fat and muscle-tone among women. Body image research has found distinct differences between the drive for thinness, the drive for muscularity and the drive for leanness (Smolak & Murnen, 2008). The drive for leanness among women refers to the motivation surrounding “having relatively low body fat and toned, physically fit muscles” (Smolak & Murnen, 2008, p.251) which logically seems to be most similar to the muscle-tone construct in the GPSCS when the operational definitions are compared. Among women, the focus of muscle-tone seems to be very similar to that of being thin with exception to the development of having muscle tone. The motivational differences in the development of a toned body are distinguishable from a thin or muscular body (Smolak & Murnen, 2008) therefore the self-perceptions of muscle-tone should also be distinct from body-fat and muscle-bulk. The accumulation of convergent validity evidence such as the negative correlation between muscle-tone and body-fat, as well as discriminant validity evidences such as the lack of correlation between muscle-bulk and body-fat aids in the interpretation of complex constructs according to theoretical rational.

According to the literature, the cultural ideal of the female body has shifted from the extremely thin body to the extremely fit and athletic but feminine body (Grogan, 2008; Homan, 2010; Steinfeldt, Carter, Benton, & Steinfeldt, 2011). Female athletes with firm toned muscles are working as models and even posing for the Sports Illustrated swimsuit issue (Gruber, 2007). Previous research examining perceptions of muscularity among women identified that women desire a toned, tight, thin and muscular body which is a hybrid of a feminine look and a strong, muscular appearance (Greenleaf, McGreer, & Parham, 2006; Homan, 2010; Markula, 1995; Steinfeldt et al., 2011). These descriptions could be referred to as a body with good muscle-tone. It has been well documented that women describe having toned muscles differently than having bulky muscles (Choi, 2003; Markula, 1995; Mosewich, Vangool, Kowalski, & McHugh, 2009). Bulky muscles are considerably larger and create a more masculine silhouette which is normally undesirable by women as they do not want to look big and manly (Mosewich, Vangool, Kowalski, & McHugh, 2009). Muscle-tone has been described as unbulky muscles, and a toned body refers to a body that is layered with long, sleek, firm, unbulky muscles (Markula, 1995). Despite the distinctions made between muscle-tone, body-fat and muscle-bulk in the literature and the subscale correlations in Study 4, EFA results did not demonstrate good discrimination between these factors within the GPSCS context.

Another aspect of the GPSCS that may affect the results of the factor analysis of the scale was the level of specificity of the instrument. The GPSCS provides a general evaluation of the muscle-tone, muscle-bulk and body-fat constructs rather than a site specific measure of PSC. Markula (1995) found that besides the size of the muscle, the location of muscle definition concerned women. Furthermore, Greenleaf et al.'s (2006) found that female aerobic exercisers and exercise instructors have body ambivalence as they are able to identify certain areas of satisfaction and areas of dissatisfaction. Furthermore, women with defined, shapely legs were more socially accepted than those having muscular arms (Greenleaf et al., 2006; Markula, 1995). For example, a participant in the Greenleaf et al. (2006) study stated she wanted "Madonna like arms" until others suggested Madonna's arms looked too "manly". In the 1980s, Jane Fonda started the womens' body sculpting movement which endorsed site specific exercises. The series of 'Jane Fonda's Workout' exercise videos lead the viewer through site-specific body toning, shaping and pulsing exercises, such as the "doggy lift" or "fire hydrant" that focused directly on problematic body areas (Markula, 1995). These site-specific toning exercises have since been highlighted in fitness magazines (Homan, 2010) and incorporated into aerobic and fitness classes and are often referred to as the torture, hard, horrible part of the class that focus on a specific problem area (Markula, 1995). The most common problem areas among women include the abdomen, thighs, underarms, and butt (Markula, 1995). More recently, fitness magazines, such as *Shape* and *Women's Fitness*, offer readers a variety of exercises that can build a tight, athletic look that is slim and lacks body-fat (Markula, 1995). The lack of distinction of the muscle-tone factor in the GPSCS may be related to the measure being at a general, overall body level rather than the assessment of specific body parts.

Reliability

The reliability of the GPSCS was assessed according to internal consistency and temporal stability (DeVellis, 2012). The internal consistency according to Cronbach's Alpha found that all four sets of items consistently correlated to their matched subscale. The temporal stability of the GPSCS according to test-retest reliability assessed how constant scores remained from one occasion to the next (DeVellis, 2012). The temporal stability of all four GPSCS subscales was good. These findings suggest that the GPSCS consistently assessed all items according to the matched construct at a single time point and over time, despite difficulty in distinguishing the structural validity of the muscle-tone construct from the muscle-bulk and body-fat constructs.

External Criterion-Related Validity

Convergent and discriminant correlations of GPSCS subscales with external variables such physical activity, exercise frequency, BMI, exercise identity, drive for muscularity, and drive for thinness supported the external construct validity of the instrument. The empirical relations between the GPSCS and these measures were demonstrated to be consistent with the meaning of the constructs in both convergent and divergent patterns.

Body mass index was positively correlated with perceptions of body-fatness and negatively correlated with perceptions of muscle-tone, as hypothesized. Unexpectedly, no relationship was found between BMI and perceptions of muscle-bulk. This lack of relationship may be due to the sample having low perceptions of muscle-bulk (M =1.78, SD = .72) resulting in a restriction in range which deflates the correlations. Furnham and Greaves (1994) found that perceptions of body size were a better indicator of body satisfaction than objective measures of body weight (BMI). This study used a self-reported measure of height and weight which was calculated into a BMI score, rather than an objective measure.

Physical activity (as measured by the GLTEQ) and exercise frequency (strength training and cardiovascular activity) were positively correlated to physical self-perceptions of muscle-tone, muscle-bulk, and muscle-strength. Previous studies have shown that physical self-concept subscales were positively correlated with self-report physical activity levels (Asci, 2005; Furnham & Greaves, 1994), however these studies did not include a self-perception of muscularity subscale. Regular exercisers report a more positive physical self-concept than non-exercisers (Tsorbatzoudis, 2005). The positive relationship between the GPSCS subscales and physical activity and exercise participation contributes to the external criterion related validity supporting the GPSCS.

Drive for muscularity refers to the “desire to achieve an idealized, muscular body type” (Morrison, Morrison, & Hopkins, 2003, p.113). The scores for the drive for muscularity were 4.90 for DMS behaviour and 4.27 for DMS attitudes on a 6-point Likert scale with a score of (1) labeled as ‘always’ and a score of (6) labeled as ‘never’. Therefore, the women generally did not have a drive for muscularity. In a previous study examining the DMS of female university students which included university student-athletes and non-athletes, the student athletes reported greater DMS than the non-student-athletes (Steinfelt, et al. 2011). Female student-athletes engaged in high-contact sports reported a greater level of DMS than athletes in low-contact sports. In the current study, behaviours related to the drive for muscularity were associated with perceptions of muscularity (tone and bulk) and muscle-strength. The observed relationship between the drive for muscularity behaviours and GPSCS was as hypothesized. The drive for muscularity attitudes were positively correlated with perceptions of muscle-tone and muscle-strength, and negatively correlated with body-fat. Unexpectedly, perception of muscle-bulk was not related to drive for muscularity attitudes which suggest that women with a greater perception of muscle bulk may not want to have large muscles. It would seem logical that women who have an undesirable amount of bulk muscles would not desire having even more muscle. A qualitative study examining female track athletes from a range of disciplines found that even female athletes have body ideals that are muscular but with a limited amount of muscle, as they do not want to be too large or bulky (Mosewich, Vangool, Kowalski, & McHugh, 2009). This rationale may account for the lack of relationship between drive for muscularity attitudes and perceptions of muscle-bulk among participants.

The women in this sample had a mean score of 3.57 for the drive for thinness on a 6-point scale where (1) suggests one always endorses a drive for thinness and (6) suggests one never endorses a drive for thinness. The drive for thinness scale was from the Eating Disorder Inventory (Garner & Olmstead, 1984) which is used to identify eating disorder symptoms in a clinical population, a strong endorsement for the drive for thinness was not expected in this University sample. The scale assesses restricting tendencies, desire to lose weight and fear of gaining weight. A positive relationship between the drive for thinness, and body satisfaction was observed among younger women (Pruis & Janowsky, 2010). The body satisfaction measure in this study included the assessment of body shape, weight and composition. In the current study, women with higher perception muscle-strength and muscle-tone scores did not endorse the drive for thinness; this was a significant inverse relationship. The drive for thinness scale is used to identify disordered eating. The pathological nature of the inferences made from the results of the drive for thinness scale, suggest a healthier level of drive for thinness from the women who perceive themselves as being strong and having muscle-tone, than those who perceive themselves as appearing to have more body-fat.

Further external criterion validity evidence was gathered from the association between exercise identity and the GPSCS subscales. A positive relationship between exercise identity, and self-perceptions of muscularity (tone and bulk) and self-perceptions of strength were found. A negative association between exercise identity and self-perceptions of body-fatness was found. Greater endorsement of a positive exercise identity includes the endorsement of self-identifying as an exerciser and beliefs that exercise is a critical component of one's identity. It seems reasonable that women identifying exercise as an important aspect of their identity also perceive themselves to be muscular and strong. In the current study, this relationship provides further criterion validity evidence supporting the GPSCS. Interestingly, muscle-tone was significantly related to every criterion measure assessed in this study (see Table 5.10). Future studies examining the relationship of the GPSCS constructs over an exercise intervention may be able to assist the understanding of exercise motivation and exercise adherence.

Limitations

This study has limitations including a purely female sample of University students. Therefore these findings cannot be generalized to men. In Study 3, the EFA among men and women demonstrated some items (Item 1, 2, 3, 5, 8, 11, 15 and 17) loading onto different factors. Among women, muscle-tone items loaded onto the muscle-bulk and body-fat factors, whereas with men the muscle-tone items loaded on to the muscle-bulk factor. Although, the female sample is a limitation, the use of women in a study examining self-perceptions of muscularity can also be viewed as a strength. The body-image literature has highlighted that body dissatisfaction concerning muscularity has been reported among men and women (Gray & Ginsberg, 2007), although there is limited work conducted among women. Muscularity has been found to be a focal point concerning body image, body dissatisfaction and drive for muscularity among men (Blashill, 2011; Davis, Karvinen, & McCreary, 2005; McCreary & Sadava, 2001; McCreary & Sasse, 2000; Olivardia, Pope, Borowiecki, & Cohane, 2004; Ridgeway & Tylka, 2005; Vartanian, Giant, & Passino, 2001); further examination in this area is needed among women.

A second limitation is the self-report assessment of height and weight, used to calculate the BMI score. According to results from the Canadian Community Health Survey, participants self-reported themselves as being approximately 0.88cm taller than their measured height and 2.33 kg less than their measure weight which results in a BMI score 1.16 kg/ m^2 lower than the actual BMI score (Elgar & Stewart. 2008).

A third limitation is the adequate but relatively small sample size for factor analysis (EFA). A sample size of over 300 cases would provide a good sample size for the use of confirmatory factor analysis (Tabachnick & Fidell, 2001, p. 628).

Summary

Body-image research attempting to capture constructs related to muscularity has primarily occurred among men (Blashill, 2011). Instruments developed to assess muscularity include the Male Body Attitudes Scale (Tylka, Bergeron, & Schwartz, 2005), the Masculine Body Ideal Distress Scale (Kimmel & Mahalik, 2004) and the Drive for Muscularity Scale (McCreary & Sasse, 2000). These instruments were specifically designed to assess body- image concerns of muscularity among men, did not include a self-perception of muscularity scale, and did not make a distinction between muscle-tone and muscle-bulk. Currently there is a lack of PSC of muscularity instruments that have supportive validity evidence in the literature, especially for women.

Validity and reliability evidence accumulated from this study established initial support for the meaningfulness and interpretability of the GPSCS scores. One of the purposes of this dissertation was to develop an instrument to assess general physical self-perceptions of muscle-tone, muscle-bulk, muscle-strength, and body-fat. Optimally, the GPSCS may be used to understand the relationship of general perceptions of muscle-tone, muscle-bulk, muscle-strength and body-fat in relation to exercise motivation, behavioural intentions to exercise and exercise behaviour. Inconsistent evidence supporting the validity and reliability for the GPSCS has been gathered. The structural validity evidence supported three factors when items measuring all four constructs namely, muscle-tone, muscle-bulk, muscle-strength and body-fat, were included in the analyzed. The muscle-tone items cannot be distinguished from the muscle-bulk and body-fat factors among women when examined with EFA. However, when muscle-tone items were assessed independently of muscle-bulk, muscle-strength and body-fat items a strong, clean single factor was observed. Furthermore, upon the examination of the inter-subscale bivariate correlations, muscle-tone and muscle-bulk had different relationships with body-fat. Muscle-tone had a significant inverse relationship with body-fat while muscle-bulk had no relationship with body-fat.

This provides construct validity evidence that muscle-tone and muscle-bulk appear to be tapping into different constructs.

The methods used to support the retention of the muscle-tone items were non-conventional among psychometricians, as muscle-tone items loaded onto two different factors among the GPSCS items. Normally these items would be removed and dropped from any further analyses. However, in this dissertation, these items were retained due to the combination of:

1. the strong single-factor solution for the muscle-tone items in an EFA (conducted independently from the other GPSCS items),

2. the different inter-subscale correlations, and

3. the external construct validity evidence supported the retention of the muscle-tone items as an independent subscale and instrument than the rest of the GPSCS items. Further studies using cognitive interviewing techniques to understand what the respondent is thinking while completing the items in the GPSCS; assessing physical self-perceptions of exercisers and athletes; assessing GPSCS scores over time, across various group settings and in response to experimental treatment and manipulations are critical in further developing the instrument and to broaden our understanding of physical self-concept (Messick, 1989).

References

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*, 131-142. doi:10 .1177/0013164485451012

Anderson, D. F., & Cychosz, C. M. (1994). Development of an exercise identity scale. *Perceptual and Motor Skills, 78*, 747-751. doi:10.2466/pms.1994. 78.3.747

Anderson, D. F., Cychosz, C. M., & Franke, W. D. (2001). Preliminary exercise identity scale (EIS) norms for three adult samples. *Journal of Sport Behavior, 24*(1), 1-9.

Asci, F. H. (2005). The construct validity of two physical self-concept measures: An example from Turkey. *Psychology of Sport and Exercise, 6*, 659-669. doi:10.1016/j.psychsport.2004.07.005

Blashill, A. J. (2011). Gender roles, eating pathology, and body image, and body dissatisfaction in men: A meta-analysis. *Body Image, 8*, 1-11. doi:10.101 6/j.bodyim.2010.09.002

Cafri, G., & Thompson, J. K. (2004). Evaluating the Convergence of Muscle Appearance Attitude Measures. *Assessment, 11*, 224-229. doi:10.1177/1 073191104267652

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81- 105. Retrieved from http://www.apa.org/pubs/journals/bul/index.aspx

Cattell, R. (1966). *Handbook of multivariate experimental psychology*. Chicago, IL: Rand McNally.

Cattell, R. (1978). *The scientific use of factor analysis in the behavioral and life sciences.* New York, NY: Plenum Press.

Cheung, M.W.L., & Yip, P.S.F. (2005). Theta reliability. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement (pp.* 791-796). Oxford, GB: Elsevier Academic Press.

Choi, P.Y.L. (2003). Muscle matters: Maintaining visible differences between women and men. *Sexualities, Evolution and Gender, 5*, 71-81. doi:10.1080/14616660310001632554

Comery, A. L. (1988). Factor analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology, 56*, 754-761. doi:10.1037//0022-006X.56.5.754

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and application. *Journal of Applied Psychology, 78*, 98-104. doi:10.1037//0021-9010.78.1.98

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 35*, 297-334. doi:10.1007/BF02310555

Davis, C., Karvinen, K., & McCreary, D. R. (2005). Personality correlates of a drive for muscularity in young men. *Personality and Individual Differences, 39*, 349-359. doi:10.1016/j.paid.2005.01.013

DeVellis, R. F. (1991). *Scale development: Theory and application*. London, GB: Sage.

DeVellis, R. F. (2012). *Scale development: Theory and application (3rd ed.).* London, GB: Sage.

Elgar, F. J., & Stewart, J. M. (2008). Validity of self-report screening for overweight and obesity. Evidence from the Community Health Survey. *Canadian Journal of Public Health, 99*, 423-427. Retrieved from www.cpha.ca

Ferguson, E., & Cox, T. (1993). Exploratory factor analysis: A user's guide. *International Journal of Selection and Assessment, 1*, 84-94. doi:10.1111/j.1468-2389.1993.tb00092.x

Furnham, A., & Greaves, N. (1994). Gender and locus of control correlates of body image dissatisfaction. *European Journal of Personality, 8*, 183-200. doi:10.1002/per.2410080304

Garner, D. M., & Olmstead, M. P. (1984). *The Eating Disorder Inventory manual.* Odessa, FL: Psychological Assessment Resources.

Godin, G., & Shephard, R. J. (1997). Godin leisure time exercise questionnaire. *Medicine and Science in Sport and Exercise, 29*(6), S36-S38. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Gorsuch, R. L. (1983). *Factor analysis (2nd ed.).* Hillsdale, NJ: Erlbaum.

Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment, 68*, 532-560. doi:10.1207/s15327752jpa6803_5

Gray, J., & Ginsberg, R. (2007). An overview of psychological and cultural research and theory. In J. K. Thompson & G. Cafri (Eds.), *The muscular ideal: Psychological, social and medical perspectives* (pp. 14-42). doi:10. 1037/11581-001

Greenleaf, C., McGreer, R., & Parham, H. (2006). Physique attitudes and self-presentational concerns: Exploratory interviews with female group aerobic exercisers and instructors. *Sex Roles, 54*, 189-199. doi:10.1007 /s11199-0 06-9337-4

Grogan, S. (2008). *Body image: Understanding body satisfaction in men, women, and children*. New York, NY: Routledge.

Homan, K. (2010). Athletic-ideal and this-ideal internalization as prospective predictors of body dissatisfaction, dieting, and compulsive exercise. *Body Image*, *7*, 240-245. doi:10.1016/j.bodyim.2010.02.004

Kimmel, S. B., & Mahalik, J. R. (2004). Measure masculine body ideal distress: Development of a measure. *International Journal of Men's Health, 3*, 1-10. doi:10.3149/jmh.0301.1

Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design*. New York, NY: Methuen.

Lautenschlager, G. J. (1989). A comparison of alternatives to conducting Monte Carlo analyses for determining parallel analysis criteria. *Multivariate Behavioral Research, 24*, 365-395. doi:10.1207/s15327906mbr2403_6

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Report, 3*, 635-694. Retrieved from http://ammons.ammons scientific.com/php/toc.php

Markula, P. (1995). Firm but shapely, fit but sexy, strong but thin: The postmodern aerobicized female body. *Sociology of Sport Journal, 12*(4), 424-453. Retrieved from http://journals.humankinetics.com/ssj

McCreary, D. R. (2007). The drive for muscularity scale: Description, psychometrics, and research findings. In J.K. Thompson & G. Cafri (Eds.), *The muscular ideal: Psychological, social, and medical perspective* (pp. 87-106). doi:10.1037/11581-004

McCreary, D. R., & Sadava, S. W. (2001). Gender differences in relationships among perceived attractiveness, life satisfaction, and health in adults as a function of body mass index and perceived weight. *Psychology of Men and Masculinity, 2*, 108-116. doi:10.1037/1524-9220.2.2.108

McCreary, D.R., & Sasse, D.K. (2000). An exploration of the drive for muscularity in adolescent boys and girls. *Journal of American College Health, 48*, 297-304. doi:10.1080/07448480009596271

McCreary, D. R., Saucier, D. M., & Courtenay, W. H. (2005). The drive for muscularity and masculinity: Testing the associations among gender-role traits, behaviors, attitudes, and conflict. *Psychology of Men and Masculinity, 6,* 83-94. doi:10.1037/1524-9220.6.2.83

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

Messick, S. (1995a). Validity of psychological assessment: Validation of inferences from persons' responses and performance as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749. doi:10.1037//0003 -066X.50.9.741

Messick, S. (1995b). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice, 14,* 5-8. doi:10.1111/j.1745-3992.1995.tb00881.x

Morrison, T. G., Morrison, M. A., & Hopkins, C. (2003). Striving for bodily perfection? An exploration of the drive for muscularity in Canadian men. *Psychology of Men and Masculinity, 4,* 111-120. doi:10.1037/1524-9220.4.2.111

Mosewich, A. D., Vangool, A. B., Kowalski, K. C., & McHugh, T. F. (2009). Exploring women track and field athletes' meanings of muscularity. *Journal of Applied Sport Psychology, 21*, 99-115. doi:10.1080/10413 200802575742

Nunnally, J. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Olivardia, R., Pope, H. G., Borowiecki, J. J., & Cohane, G. H. (2004). Biceps and body image: The relationship between muscularity and self-esteem, depression, and eating disorder symptoms. *Psychology of Men and Masculinity, 5,* 112-120. doi:10.1037/1524-9220.5.2.112

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach.* Hillsdale, N J: Erlbaum.

Pruis, T. A., & Janowsky, J. S. (2010). Assessment of body image in younger and older women. *The Journal of General Psychology, 137*, 225-238. doi:10. 1080/00221309.2010.484446

Ridgeway, R. T., & Tylka, T. L. (2005). College men's perceptions of ideal body composition and shape. *Psychology of Men and Masculinity, 6*, 209-220. doi:10.1037/1524-9220.6.3.209

Smolak, L., & Murnen, S. K. (2008). Drive for leaness: Assessment and relationship to gender role and objectification. *Body Image, 5*, 251-260. doi:10.1016/j.bodyim.2008.03.004

Statistics Canada. (2008). *Appendix C: Comparison of ethnic origins disseminated in 2006, 2001 and 1996*. Retrieved from http://www12. statcan.ca/census-recensement/2006/ref/dict/app-ann003-eng.cfm

Steinfeldt, J. A., Carter, H., Benton, E., & Steinfeldt, M. C. (2011). Muscularity beliefs of female college student-athletes. *Sex Roles, 64*, 543-554. doi:10 .1007/s11199-011-9935-2

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics (5th ed.).* Toronto, ON: Pearson.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics (4th ed.).* Toronto, ON: Pearson.

Thomas, J. R., Nelson, J. K., & Silverman, S. J. (2005). *Research methods in physical activity* (5th ed.). Champaign, IL: Human Kinetics.

Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago, IL: University of Chicago Press.

Trochim, W. M. K. (2006). *Research methods knowledge base*. Retrieved from http://www.socialresearchmethods.net/kb/index.php

Tsorbatzoudis, H. (2005**).** Psychometric evaluation of the Greek Physical Self-Description Questionnaire. *Perceptual and Motor Skills, 101*, 79-89. doi:10.2466/pms.102.3.703-718

Tylka, T. L., Bergeron, D., & Schwartz, J. P. (2005). Development and psychometric evaluation of the Male Body Attitude Scale (MBAS). *Body Image, 2*, 161-175. doi:10.1016/j.bodyim.2005.03.001

Vartanian, L. R., Giant, C. L., & Passino, R. M. (2001). “Ally McBeal vs. Arnold Schwarzenegger”: Comparing mass media, interpersonal feedback and gender as predictors of satisfaction with body thinness and muscularity. *Social Behavior and Personality, 29*(7), 711-724. Retrieved from http://www.sbp-journal.com/index.php/sbp/index

Wilson, P. M., & Muon, S. (2008). Psychometric properties of the exercise identity scale in a university sample. *Exercise Identity Measurement, 6*, 115-131. doi:10.1080/1612197X.2008.9671857

Zabinski, M. F., Calfas, K. J., Gehrman, C. A., Wilfley, D. E., & Sallis, J. F. (2001). Effects of activity intervention on body image. *Annals of Behavioral Medicine, 23*, 247-252. doi:10.1207/S15324796ABM2304_3

General Discussion / Conclusions

Physical self-concept is an important social cognitive factor with a unique position in the self-system due to the body's abilities, attributes and appearance (Fox, 2000). Associations between physical self-concept and health behaviours, such as physical activity and dietary behaviours, and emotions, such as anxiety and lower self-esteem, support the importance of PSC (Crocker et al., 2001; Crocker et al., 2003; Dunton, Jamner, & Cooper, 2003; Fox, 2000; Leary, 1992). Interestingly, physical self-conceptions of the body's appearance are more strongly related to global self-esteem than perceptions of physical ability or health (Crocker, Sabiston, Kowalski, McDonough, & Kowalski, 2006; Harter, 1999). Body self-concept has traditionally focused on body-fatness (Fox & Corbin, 1989; Marsh, 1996). Over the last 20 years, ideal body composition for men and women has become more athletic and muscular which has resulted in experiences of body-dissatisfaction in regard to muscularity (Gray & Ginsberg, 2007; Markula, 1995; Thompson & Cafri, 2007). The most commonly used quantitative PSC instruments (PPSP: Fox, 1990; PSDQ: Marsh, Richards, Johnson, Roche, & Tremayne, 1994) lack a self-perception of muscularity scale; even though qualitative body-image literature has identified muscularity as an important aspect of body ideals and PSC among men and women (Greenleaf, McGreer, & Parham, 2006; Markula, 1995; Mosewich, Vangool, Kowalski, & McHugh, 2009; Ridgeway & Tylka, 2005). Given this omission, it seemed reasonable to develop an instrument that can assess physical self-concept of muscularity. Therefore, the purposes of this dissertation were to develop an instrument that assesses the physical self-concept of muscularity from a general, overall body perspective and provide evidence of validity and reliability for the new instrument.

To achieve this purpose, the GPSCS was developed as a measure of self-perceptions of muscle-tone, muscle-bulk, muscle-strength, and body-fat. Evidence of the validity and reliability of the items representing the constructs of muscle-bulk, muscle-strength and body-fat in the GPSCS was developed over a series of four studies in this dissertation. Although the muscle-tone items were not distinguishable from self-perceptions of muscle-bulk and body-fat using exploratory factor analysis, content relevance and representativeness validity evidence supported these items. Furthermore, the muscle-tone construct was correlated as expected to every criterion measure (drive for muscularity, drive for thinness, exercise identity, BMI, exercise engagement, and leisure time physical activity) assessed in study 4 which highlights the importance of the construct to other body related constructs and provides support for the continued consideration of muscle-tone. It is proposed, therefore, that the items representing the muscle-tone construct may be used as a separate instrument to assess self-perceptions of muscle-tone.

The findings from the series of studies brought forward some limitations, concerns, discussion points, and ideas for future directions of the GPSCS. Some areas worthy of exploration include: 1) examining the distinctness between the muscle tone, muscle bulk and body fat constructs; 2) exploring the level of specificity of PSC measurement, for example a general measure of muscle-tone versus a body-site specific measure of muscle-tone; 3) discussing the development of physical self-concept; 4) examining issues associated with sex and the GPSCS.

Distinctions between the muscle tone, muscle bulk and body fat constructs

According to the literature muscle-tone and muscle-bulk are two distinct constructs that have different ideals according to sex (Markula, 1995; Ridgeway & Tylka, 2005), and sexual orientation among men (Yelland & Tiggemann, 2003). The ideal female body and homosexual male body is a lean, slender and toned body with visible muscles which is considered a toned body (Choi, 2000; Gruber, 2007; Markula, 1995; McCabe, Ricciardelli, & Finemore, 2002; Yelland & Tiggemann, 2003). Muscle-tone takes into consideration the development of muscle as well as the amount of body-fat. In order for a muscle to be defined or “cut”, muscle bulk needs to be present and fat needs to be at a minimum level. Muscle-bulk is the term used to describe a large amount of muscle mass and does not take into consideration body-fat. In general, muscle-bulk is desired by men and considered unattractive among women. The series of dissertation studies vary in the levels of support for muscle-tone, muscle-bulk and body-fat as psychometrically distinct constructs.

The content validity studies in this dissertation (study 1 and study 2) support the muscle-tone, muscle-bulk and body-fat constructs. The quantitative aspect of these studies identified the degree of match between the GPSCS items and constructs, and representativeness of the set of GPSCS items but fails to offer any explanation or solution for the fair scores. Essentially, these two studies identified the degree of match each item had according to the construct of interest and assessed the ability of the groups of items to represent each of the constructs. An item with a fair match to the construct of interest is a problematic item. These findings suggested that for the most part, the items matched the construct of interest. However, seven items were scored as a fair match to the construct of interest in study 1 and one item was scored as a fair match according to the construct of interest in study 2. The qualitative data was critical in identifying how and why some items could be improved to reflect the construct from a theoretical and practical perspective (i.e., identification of different terms, comments on why the key descriptor in the item is or is not appropriate). The findings from these studies supported the distinction of muscle-tone, muscle-bulk, and body-fat, as identified in the literature (Choi, 2000; Gruber, 2007; Markula, 1995; McCabe, Ricciardelli, & Finemore, 2002; Ridgeway & Tylka, 2005; Yelland & Tiggemann, 2003).

Some of the greatest support for the GPSCS constructs came from the qualitative feedback from the fitness experts in study 2. Some examples of support for the distinction between the GPSCS muscle-tone, muscle-bulk, and body-fat, constructs included the terms used to express each of these constructs by fitness professionals and their clients. The muscle-tone construct had the longest list of terms (i.e., lean, long, not bulky, juicy, sculpted, defined, avoid big muscles, muscles with the removal of the excess flab, etc.). The substantial amount of terms for muscle-tone highlights the prominence of the idea in the fitness setting. Although the fitness experts were not asked to determine if the constructs were distinct, the experts' feedback provided evidence that muscle-tone was not the same as muscle-bulk. Muscle-tone was described as the presence of smaller muscles and an absence of fat or 'jiggling', being able to see the shape of the muscle without being bulky. Muscle-bulk was described by the fitness experts as having huge, large, and big, muscles. These terms for muscle-tone and muscle-bulk are similar to those found in the muscularity literature (Choi, 2000; Gruber, 2007; Markula, 1995; McCabe, Ricciardelli, & Finemore, 2002; Ridgeway & Tylka, 2005; Yelland & Tiggemann, 2003). Furthermore, the descriptors of body-fat from the qualitative feedback included having an 'extra tire' around the middle, flabby, soft, excess, overweight, etc. These terms are distinct from the muscle-tone and muscle-bulk terms. The combination of quantitative content relevance and representativeness results with the qualitative feedback supports the match between items and constructs.

The findings from study 3 and study 4 did not provide clear structural validity evidence for the GPSCS according to the exploratory factor analysis results. Exploratory factor analysis basically determines the number of latent variables in a set of items (Brown, 2006). When the complete set of GPSCS items was included in the EFA, the muscle-tone items did not unequivocally fall on a single factor, but also fell onto the muscle-bulk and body-fat factors. Upon re-assessment of the GPSCS excluding the muscle-tone items, a clean simple structure was observed which suggested three distinct constructs (muscle-bulk, body-fat and muscle-strength). When assessed separately, the muscle-tone items comprised a strong single factor. These results suggest that muscle-tone cannot be distinguished statistically from muscle-bulk and body-fat despite the support within the first two studies and in the body image literature for the existence of the idea. Factor analysis is one method of guiding the decision making process regarding the differentiation of latent constructs; common sense should also be used in combination with these analytical techniques when using the factor analysis evidence to make decisions (DeVillis, 2012, p. 153).

The purpose of EFA is to identify the most parsimonious number of latent variables that account for variation and covariation of observable indicators to determine interpretable factors (Brown, 2006). Although achieving a simple pattern structure is an important aspect of factor analysis, examining different analytical techniques to "let the data speak" is a critical aspect of the interpretability of the solution (Thompson, 2004). The unfavorable EFA findings were not able to distinguish between the GPSCS constructs. According to a comprehensive construct validity approach, the integration of all evidence that 'bears on the meaning of the test scores' are important aspects of construct validity (Messick, 1995a, p.742). Therefore the EFA results should be taken into consideration with the entire set of evidence for the GPSCS constructs from analyses in studies 1, 2, 3 and 4. According to Messick (1995a), validity is an evolving property that accounts for "the overall evaluation of the degree to which empirical evidence and theoretical rationales support the adequate and appropriateness of interpretations and action on the basis of test scores" (p. 741).

The muscle-tone construct includes having a low degree of body-fat and having some degree of muscle bulk. This combination of attributes may be the challenging aspect of the assessment of muscle-tone. Although the content validity and criterion validity studies have provided strong evidence supporting the GPSCS, the findings from the structural validity studies have not been optimal. Future studies using other analytical techniques, such as confirmatory factor analysis (CFA), item response theory, and multidimensional scaling, may provide further evidence to guide the interpretability of the muscle-tone construct..Exploratory factor analysis identifies the group of items that ‘hang together’ according to an underlying latent variable; EFA allows all items to load onto all factors. Confirmatory factor analysis assesses an a priori hypothesized model which suggests that specific items are indicators of specific factors. Both EFA and CFA are analytical techniques that utilize the respondents’ self-evaluation of the item content (i.e., response to the GPSCS items) for the analysis. Multidimensional scaling is unique, as the respondents rate the degree of similarity between each set of items rather than the response to the GPSCS items.

Exploring the level of specificity of PSC measurement

One explanation for the inconsistent findings, between the content relevance and representativeness studies and the factor analysis studies, may be attributable to the level of specificity of the GPSCS constructs and/or items. The content relevance and representativeness studies supported the GPSCS items and constructs as matching, and representing the operational definitions for muscle-tone, muscle-bulk, muscle-strength, and body-fat. In addition to these findings, the fitness experts in study 2 identified that muscle-tone, muscle-bulk, and body-fat concerns were often discussed according to specific body-sites. For example, fitness clients commonly discussed concerns with having jiggly ‘bat wings’, an 'extra tire around the middle', bulky thighs, and often pointed to body-sites that were, in their opinions, too fat or lacked muscle tone. Additionally, some clients described their body ideals as having tight thighs, or a beach body with big biceps, shoulders and chest. These comments highlight the spontaneous use of body-site specific self-assessments which suggest that the GPSCS constructs may be more appropriate as body-site specific evaluations. Previously, qualitative studies have identified that men and women hold different site specific body ideals (Andersen et al., 2000; Furnham et al., 2002; Markula, 1995; Rideway & Tylka, 2005). More specifically, women focus on their abdomen, thighs, upper arms, and butt (Furnham et al., 2002; Markula, 1995), whereas, men focus on their chest, arms, abdomen, and shoulders (McFarland & Petrie, 2012; Rideway & Tylka, 2005). These studies provide further evidence suggesting that muscle-tone and muscle-bulk are PSC constructs that people reflect upon according to body-site rather than general body self-perceptions. The body-site specific body image literature may provide insight and possible explanation for the difficulty in distinguishing muscle-tone and muscle-bulk and body fat as distinct factors within the GPSCS. It may not be the conceptualization of muscle-tone that is problematic, rather the lack of specificity in the items within the measurement instrument. Further exploration of a body-site specific measurement of the PSC constructs of muscle-tone, muscle-bulk, and body-fat would be a valuable future direction.

Development of physical self-concept

According to Harter (1999), the self is both cognitively and socially constructed. Cognitive construction refers to the use of theories individuals make about the world, the self, and their experiences which alter the structure of the self-system (Harter, 1999). “A person’s perceptions, thoughts, and feelings about his or her body” (Schilder, 1950, p.11) are subjective and, both cognitively and socially determined (Grogan, 1999). Factors influencing the self include the actual self-concept structure, the impact of reactions to socialization, and social context (Harter, 1999).

The third study of this dissertation found that PSC scores were significantly different among men and women. The women demonstrated poorer perceptions of muscle-tone, muscle-bulk, muscle-strength, and body-fat. These findings are consistent with previous PSC studies (Grogan, 1999) and provide construct validity evidence supporting the use of the GPSCS. According to the PSC literature, physical self-perceptions are influenced by sex, sexual orientation, age, subculture membership, and socioeconomic status (Cash & Pruzinsky, 2002). For example, women have poorer physical self-concepts than men (Grogan, 1999), whereas, homosexual men have poorer PSCs than heterosexual men (Yelland & Tiggemann, 2003). The relationship between sociodemographic factors and PSC identifies that certain sociodemographic group memberships (i.e., Caucasian female adolescents) are related to poorer levels of PSC.

The findings from the fourth study of this dissertation established that women with higher BMIs had significantly poorer perceptions of body-fatness and muscle-tone than women with lower BMIs. This would suggest that actual body weight (BMI) impacts PSC among University-aged women. According to the literature, physical self-perceptions are not completely determined by actual body shape, composition or structure, rather by the self-structure and the self-evaluations within the social context (Furnham & Greaves, 1994; Grogan, 1999, p.166). Factors such as social experiences have been found to be better predictors of PSC levels than actual body size (Grogan, 1999). Self-evaluations within the social context of being a University-aged woman may be the reason why women with higher BMIs have poorer PSC, but these relationships were not tested in this study.

Examining issues associated with sex and the GPSCS

The findings from study 3 identified that men and women responded to the GPSCS differently. Women scored significantly lower than men in all subscales except body-fat on which they were significantly higher. This highlights the notion that sex specific instruments may be more appropriate for the assessment of PSC than a uni-sex instrument. Men and women have different body ideals (Markula, 1995; Ridgeway & Tylka, 2005) and have different perceptions of desirable PSC scores. Both men and women are concerned with their level of body-fat, muscle-tone, muscle-bulk and muscle-strength, although the ideal levels of these concerns seems to differ according to sex. For example, among women a high score in muscle-bulk would be undesirable, whereas men may consider a high score as appealing. Furthermore, it is possible that there are body-site specific concerns that also distinguish people based on sex, gender and sexual orientation that might be best addressed separately based on these socio-demographic factors. Although this was not assessed in this dissertation this would be an important direction for future research.

Regardless of any sex differences in body ideals, the operational definition of the PSC constructs remains constant. Content experts in study 1 and study 2 identified sex-specific terms used in the GPSCS items, such as plus sized and slender. One option in instrument development is to embrace the sex-specific terminology and include the language used by men and women to develop sex-specific instruments to enhance the ecological validity of the measure. This instrument could be further specified according to culture, as PSC terminology used among different cultural groups differs as well (i.e., the term for muscle-tone in the United Kingdom is 'henched'). To manage these concerns in this dissertation, the GPSCS items were modified and sex-specific terms were eliminated. The target population for the instrument was narrowed to assess North American University aged men and women. A future direction of this dissertation would be to possibly develop PSC instruments that target people based on sex, gender, sexual orientation, and culture. The ability of the GPSCS to assess heterogeneous scores among men and women in study 3 demonstrates the sensitivity of the instrument to detect PSC differences according to sex, this is one form of construct validity. Further validity evidence explaining the sex specific aspects of PSC may be established by examining the structural validity of the GPSCS using confirmatory factor analysis and multidimensional scaling according to sex, as well as, criterion validity studies assessing similarities and differences in responses to the GPSCS according to sex and other socio-demographic factors. Once more, the accumulation of different forms of evidence supporting the use of the GPSCS for the assessment of PSC is required to further develop validity evidence.

Strengths of the Dissertation

Types of validity evidence

According to Messick (1995a), the six criteria of construct validity psychological measurement include content, substantive, structural, generalizability, external and consequential aspects. This dissertation provided validity evidence for three out of the six aspects of construct validity.

Content validity. The content validity aspect of this dissertation included the assessment of content relevance, representativeness and technical qualities of the questionnaire items (Messick, 1989) by two different groups of experts in study 1 and 2. These studies focused on the specifications and boundaries of the PSC constructs to insure all important aspects of the construct domain were covered (Messick, 1995a). The first study included the traditional appraisal by academic experts, whereas the second study included professional experts from the fitness industry. The combination of psychometric training and practical experience of this set of experts was one of the greatest strengths of this dissertation. This combination of experts provided rich qualitative data that was used to improve the GPSCS and further our understanding about the PSC.

Structural validity. Structural validity refers to the fidelity of the scoring structure to the construct domain (Messick, 1989). The internal structure of the

GPSCS was assessed in two studies. Factor analysis is a common form of developing structural validity evidence. Optimally, factor analysis requires a sample size of greater than 200 respondents. This criterion was achieved in the sample of women in studies 3 and 4, but not among the male sample. Therefore, having a large sample size among women in this dissertation was a strength of the procedures and findings in the disseration.

External validity. The external aspect of validity in this dissertation included criterion related evidence (Messick, 1995). This refers to the extent to which the hypothesized relationships between the GPSCS and external measures behave as predicted. The fourth study in the dissertation contained multiple criterion measures which provided evidence of the utility of the scores for an applied purpose (Messick, 1995). This is a third source of validity evidence from the dissertation. The combination of content, structural and external validity evidence from these initial instrument development studies was a strength of this dissertation.

Qualitative data.

Although this dissertation focused on instrument development using quantitative approaches, the qualitative aspects of the dissertation resulted in substantial improvement to the GPSCS. Qualitative approaches can explore the meaning of the items and constructs beyond those currently used in quantitative approaches (Bartunek & Seo, 2002). The qualitative data assisted with the rewording of items, and identified areas of concern and strength in the instrument. More specifically, the experts provided rich qualitative feedback, which included common terms used among clients to describe muscle-tone, muscle-bulk, muscle-strength, and body-fat. Terms such as ripped, well-sculpted, jiggles and flabby stemmed from the fitness expert feedback. Furthermore, having an international expert panel in study 1 identified the inability of the terms to be used universally across cultures which lead to the narrowing of the scope of the instrument to target North Americans. This group was very useful and important in the development of new items as well as identifying terms that were problematic according to personal experiences and conversations with the general public.

Limitations

Some limitations of these studies included the modest number of men in study 3 and lack of men in study 4. Increasing the representation of men in the samples and increased sample size overall would increase the stability of the factor analysis, and allow for assessment of parallel structure between socio-demographic factors. This is one method to improve upon these dissertation studies.

A second limitation of study 3 and study 4 was the recruitment of the samples from University classes. Students from the University community may respond to the surveys differently than people from the general population, which reduces the generalizability of the instrument across populations. Increasing the scope of population representativeness to include different types of exercisers, such as, fitness centre members, an age range of participant (18 years to 35 years of age), people of different weight status (under-weight, normal weight, overweight and obese), and people from various levels of socioeconomic status would increase the capability to assess the generalizability of the instrument.

Future directions

The findings suggest that the GPSCS is a promising instrument for the assessment of muscle-tone, muscle-bulk, muscle-strength, and body-fat. The process of construct validation is ongoing (Messick, 1995). Future research may include cognitive interviews with participants that have completed the GPSCS. Cognitive interviewing focuses on the cognitive processes that respondents use to answer items (Willis, DeMaio, & Harris-Kojetin, 1999). A combination of the think-aloud and verbal probing techniques could be used to further understand cognitive processes in GPSCS item responses. The qualitative data obtained from cognitive interviewing may provide direction concerning the difficulty in the measurement of muscle-tone and muscle-bulk lies. More specifically, this may provide evidence of support for a general versus body site-specific PSC instrument depending on how the participants determine their current responses to the GPSCS items. Furthermore, this may provide some insight into the different processes men and women use to interpret the items on the GPSCS.

Multidimensional scaling (MDS) is a quantitative approach that may possibly be used to further the understanding of the latent dimensionality of the self-perception of muscle-tone and muscle-bulk items (Dunn, 1999). In MDS studies, participants would examine each item pairing combination from all the muscle-tone and muscle-bulk items and assess the pair on the degree of similarity.

These similarity rating assessments would assist in the clarification of the latent structure of the muscle-tone and muscle-bulk items (Dunn, 1999). Therefore, this would provide evidence supporting or dismissing the distinctness of the GPSCS items to their matched construct and provide evidence for the distinct GPSCS constructs.

Confirmatory factor analysis (CFA) of the GPSCS among University aged males and females is a critical next step in the development of validity evidence. The content relevance, content representativeness and criterion-related validity evidence provides support for the theoretical rationale supporting the proposed structure of the GPSCS. Separate CFAs for men and women should be conducted to test the consistency of the correlations among variables according to the hypothesized factor structure (Tabachnick & Fidell, 2007, p. 610). This should be followed by a second CFA in a separate sample, prior to being assessed in other population subgroups (Brown, 2006). Currently, the EFA was not able to demonstrate four distinct constructs from the GPSCS; this may be a limitation of EFA or it may identify a problem with the GPSCS items. Conducting a CFA would provide evidence regarding the causes of the GPSCS issues.

Further evidence of construct validity should be developed by assessing the expected performance of the GPSCS constructs between subgroups. For example, assessing the similarities and differences of GPSCS scores among men and women (Grogan, 1999), homosexual and heterosexual males (Yelland & Tiggemann, 2003), exercisers and non-exercisers, different type of athletes (i.e., long distance runner versus sprinter, bodybuilder vs. dancer) (Mosewich et al.,

2009) and people in different age categories. A match between the empirical evaluation and theoretical rationale underlying score interpretation would provide a strong degree of evidence supporting the GPSCS (Messick, 1995a). Group membership provides both convergent evidence supporting the construct and discriminant evidence discounting plausible rival interpretations which contributes to the construct validity of psychological assessment (Messick, 1995b).

Additional construct validity evidence could be developed by assessing the GPSCS over time, and in response to treatment such as body image counseling or participation in a body image workshop (Messick, 1995b). Evaluating the GPSCS constructs over the initiation and maintenance of an exercise program and/or healthy eating program may provide validity evidence for the instrument and possibly provide further information about the relationship between PSC and health behaviour change. Assessing criterion-related constructs that would be related to the GPSCS constructs as one aspect of the intervention would provide validity evidence and further understanding of the relationship of the PSC in exercise initiation and maintenance.

Another possible future direction is to expand the PSC literature by examining the use of a body-site specific instrument to assess muscle-tone and muscle-bulk. Recently, McFarland and Petrie (2012) developed a site-specific body satisfaction scale for men that assessed the face, upper-body and lower-body. A similar instrument could be developed to assess PSC. Increasing the specificity of the instrument to include body-sites that are relevant to men and women and their ideals may enhance our understanding of the role of the physical self-concept of muscle-tone and muscle-bulk.

References

Anderson, D. F., Cychosz, C. M., & Franke, W. D. (2001). Preliminary exercise identity scale (EIS) norms for three adult samples. *Journal of Sport Behavior*, *24*, 1-9.

Bartunek, J.M., & Seo, M. (2002). Qualitative research can add new meaning to quantitative research. *Journal of Organizational Behaviour*, *23*, 237-242.

Boisvert, J. A., & Harrell, W. A. (2009). Homosexuality as a risk factor for eating disorder symptomatology in men. *The Journal of Men's Studies, 17*, 210-225. doi:10.3149/jms.1703.210

Boroughs, M., & Thompson, J. K. (2002). Exercise status and sexual orientation as moderators of body image disturbance and eating disorders in males. *International Journal of Eating Disorders, 31*, 307-311. doi:10.1002/eat.10031

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Brown, T. A., Cash, T. F., & Milkulka, P. J. (1990). Attitudinal body-image assessment: factor analysis of the body self-relations questionnaire. *Journal of Personality Assessment, 55*, 135-144. doi:10.1080/00223891.1990.9674052

Cattell, R. (1978). *The scientific use of factor analysis in the behavioral and life sciences.* New York, NY: Plenum Press.

Cash, T. F., & Pruzinsky, T. (2002). *Body image: A handbook of theory, research, and clinical practice.* New York, NY: Guildford Press.

Choi, P. Y. L. (2003). Muscle matters: Maintaining visible differences between women and men. *Sexualities, Evolution and Gender, 5*, 71-81. doi:10.1080/14616660310001632554

Comrey, A. (1978). Common methodological problems in factor analytical studies. *Journal of Consulting and Clinical Psychology, 46*, 648-659. doi:10.1037/0022-006X.46.4.648

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis.* Hillsdale, NJ: Erlbaum.

Crocker, P. R. E., Kowalski, N., Kowalski, K., Chad, K., Humbert, L., & Forrester, S. (2001). Smoking behavior and dietary restraint in young adolescent women: The role of physical self-perceptions. *Canadian Journal of Public Health, 92*, 428-432. Retrieved from http://journal.cpha.ca/index.php/cjph/issue/view/171

Crocker, P. R. E., Sabiston, C. M., Forrester, S., Kowalski, N. C., Kowalski, N., & McDonough, M. (2003). Predicting changes in physical activity, dietary restraint, and physique anxiety in adolescent girls: Examining covariance in physical self-perceptions. *Canadian Journal of Public Health, 94,* 332-337. Retrieved from http://journal.cpha.ca/index.php/cjph/issue /view/183

Crocker, P. R. E., Sabiston, C. M., Kowalski, K. C., McDonough, M.H., & Kowalski, N. (2006). Longitudinal assessment of the relationship between physical self-concept and health-related behavior and emotion in adolescent girls. *Journal of Applied Sport Psychology, 18,* 185-200. doi:10.1080/10413200600830257

Dunn, J.G.H. (1999). A theoretical framework for structuring the content of competitive worry in ice hockey. *Journal of Sport and Exercise Psychology, 21*, 259-279. Retrieved from http://journals.human kinetics.com/jsep

Dunton, G. F., Jamner, M. S., & Cooper, D. M. (2003). Physical self-concept in adolescent girls: Behavioral and physiological correlates. *Research Quarterly for Exercise and Sport, 74 (3),* 360-365.

Ferguson, E., & Cox, T. (1993). Exploratory factor analysis: A user's guide. *International Journal of Selection and Assessment*, *1*, 84-94. doi:10.1111 /j.1468-2389.1993.tb00092.x

Fox, K. R. (1990). *The Physical Self-Perception Profile Manual.* DeKalb, IL: Northern Illinois University.

Fox, K. R. (2000). Self-esteem, self-presentations and exercise. *International Journal of Sport Psychology, 31*, 228-240.

Fox, K. R., & Corbin, C. B. (1989). The physical self-perception profile: Development and preliminary validation. *Journal of Sport and Exercise Psychology, 11*, 408-430. Retrieved from http://journals.humankinetics .com /jsep

Furnham, A., Badmin, N., & Sneade, I. (2002). Body image dissatisfaction: Gender differences in eating attitudes, self-esteem, and reasons for exercise. *The Journal of Psychology, 136*, 581-596. doi:10.1080/00223980209604820

Furnham, A., & Greaves, N. (1994). Gender and locus of control correlates of body image dissatisfaction. *European Journal of Personality, 8*, 183-200.

Franzoi, S. L., & Shields, S. A. (1984). The body esteem scale: Multidimensional structure and sex differences in college populations. *Journal of Personality Assessment, 48*, 173-178. doi:10.1207/s15327752jpa4802_11

Garner, D. M., Olmstead, M. A., & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders, 2*, 15-34. doi:10.1002/1098-108X(198321) 2:2<15::AID-EAT2260020203>3.0.CO;2-6

Gray, J. J., & Ginsberg, R. L. (2007). Muscle dissatisfaction: An overview of psychological and cultural research and theory. In J. K. Thompson & G. Cafri (Eds.), *The muscular ideal: Psychological, social, and medical perspectives* (pp.15-39). doi:10.1037/11581- 011

Groesz, L. M., Levine, M. P., & Murnen, S. K. (2002). The effect of experimental presentation of thin media images on body satisfaction: A meta-analytic review. *International Journal of Eating Disorders, 31*, 1-15.

Grogan, S. (1999). *Body image: Understanding body dissatisfaction in men, women and children.* New York, NY: Routledge.

Gruber, A. J. (2007). A more muscular female body ideal. In J. K. Thompson & G. Cafri (Eds.), *The muscular ideal: Psychological, social, and medical perspectives* (pp.217-234). doi:10.1037/11581-011

Gruber, A. J., Pope, H. G., Borowiecki, J., & Cohane, G. (1999). The development of the somatomorphic matrix: A bi-axial instrument for measuring body image in men and women. In T. S. Olds, J. Dollman & K. I. Norton (Eds.), Kinanthropometry VI. Sydney: international Society for the Advancement of Kinanthropometry.

Harrison, K., & Bond, B. J. (2007). Gaming magazines and the drive in preadolescent boys: A longitudinal examination. *Body Image, 4,* 269-277. doi:10.1016/j.bodyim.2007.03.003

Harter, S. (1999). *The construction of the self: A developmental perspective.* New York: NY: Guildford Press.

Jacobi, L., & Cash, T. F. (1994). In pursuit of the perfect appearance: Discrepancies among self-ideal perceptions of multiple physical attributes. *Journal of Applied Social Psychology, 24*, 379-396. doi:10.1111/j.1559-1816.1994.tb00588.x

Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design.* New York, NY: Methuen and Co.

Leary, M. (1992). Self-presentational processes in exercise and sport. *Journal of Sport and Exercise Psychology, 14,* 339-351. Retrieved from http://journals.humankinetics.com /jsep-back-issues

Leit, R. A., Gray, J. J., & Pope, H. G. J. (2002). The media's representation of the ideal male body: A cause for muscle dysmorphia?. *International Journal of Eating Disorders, 31*, 334-338. doi: 10.1002/eat.10019

Leit, R. A., Pope, H. G., & Gray, J. J. (2001). Cultural expectations of muscularity in men: The evolution of Playgirl centerfolds. *International Journal of Eating Disorders, 29*, 90-93. doi:10.1002/1098-108X(200101)29:1<90::AID-EAT15>3.0.CO;2-F

Lynch, S. M., & Zellner, D. A. (1999). Figure preferences in two generations of men: The use of figure drawings illustrating differences in muscle mass. *Sex Roles, 40*, 833-843. doi:10.1023/A:1018868904239

Markula, P. (1995). Firm but shapely, fit but sexy, strong but thin: The postmodern aerobicized female body. *Sociology of Sport Journal*, *12*, 424-453. Retrieved from http://journals.humankinetics.com/ssj

Marsh, H. W. (1996). Construct validity of Physical Self Description Questionnaire responses: Relations to external criteria. *Journal of Sport and Exercise Psychology, 18*, 111-131. Retrieved from http://journals.humankinetics.com/jsep

Marsh, H. W., Richards, G. E., Johnson, S., Roche, L., & Tremayne, P. (1994). Physical Self- Description Questionnaire: Psychometric properties and a multitrait-multimethod analysis of relations to existing instruments. *Journal of Sport and Exercise Psychology, 16*, 270-305. Retrieved from http://journals.humankinetics.com/jsep

McCabe, M.P., Ricciardelli, L.A., & Finemore, J. (2002). The role of puberty, media and popularity with peers on strategies to increase weight, decrease weight and increase muscle tone among adolescent boys and girls. Journal of Psychosomatic Research, 52, 145-153. doi:10.1016/S0022-3999(01)00272-0

McCreary, D. R., & Sasse, D. K. (2000). An exploration of the drive for muscularity in adolescent boys and girls. *Journal of American College Health, 48*, 297-304. doi:10.1080/07448480009596271

McFarland, M. B., & Petrie, T. A. (2012). Male body satisfaction: factorial and construct validity of the body part satisfaction scale for men. *Journal of Counseling Psychology, 59*, 329-337. doi:10.1037/a0026777

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York, NY: MacMillan.

Messick, S. (1995a). Validity of psychological assessment: Validation of inferences from persons' responses and performance as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749. doi:10.1037//0003-066X.50.9.741

Messick, S. (1995b). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice, 14,* 5-8. doi:10.1111/j.1745-3992.1995.tb00881.x

Mosewich, A. D., Vangool, A. B., Kowalski, K. C., & McHugh, T. F. (2009). Exploring women track and field athletes' meanings of muscularity. *Journal of Applied Sport Psychology, 21*, 99-115. doi:10.1080/104 13200802575742

Morrison, T. G., Morrison, M. A., & Hopkins, C. (2003). Striving for bodily perfection? An expectation of the drive for muscularity in Canadian men. *Psychology of Men and Muscularity, 4,* 111-120. doi:10.1037/1524-92 20.4.2.111

Olivardia, R., Pope, H. G., Borowiecki, J. J., & Cohane, G. H. (2004). Biceps and body image: The relationship between muscularity and self-esteem, depression, and eating disorder symptoms. *Psychology of Men and Masculinity, 5*, 112-120. doi:10.1037/1524-9220.5.2.112

Pope, H. G., Olivardia, R., Borowiecki, J. J., & Cohane, G. H. (2001). The growing commercial value of the male body: A longitudinal survey of advertising in women's magazines. *Psychotherapy and Psychosomatics, 70*, 189-192. doi:10.1159/000056252

Pope, H. G., Olivardia, R., Gruber, A., & Borowiecki, J. (1999). Evolving ideals of male body image as seen through action toys. *International Journal of Eating Disorders, 26*, 65-72. doi:10.1002/(SICI)1098108X(199907)26:1 <65::AID-EAT8>3.0.CO;2-D

Ridgeway, R. T., & Tylka, T. L. (2005). College men's perceptions of ideal body composition and shape. *Psychology of Men and Masculinity, 6*, 209-220. doi:10.1037/15249220.6.3. 209

Secord, P. F., & Jourard, S. M. (1953). The appraisal of body cathexis: Body cathexis and the self. *Journal of Consulting Psychology, 17*, 343-347. doi:10.1037/h0060689

Smith, A. R., Hawkeswood, S. E., Bodell, L. P., & Joiner, T. E. (2011). Muscularity versus leanness: An examination of body ideals and predictors of disordered eating in herterosexual and gay college students. *Body Image, 8*, 232-236. doi:10.1016/j.bodyim. 2011.03.005

Spitzer, B. L., Henderson, K. A., & Zivian, M. T. (1999). Gender differences in population versus media body size: A comparison over four decades. *Sex Roles, 40*, 545-565. doi:10.1023/A:1018836029738

Stunkard, A. J., Sorensen, T. & Schulsinger, F. (1983). Use of the Danish adoption registrar for the study of obesity and thinness. In S. S. Kety, L. P. Rowland, R. L. Sidman, & S. W. Matthysse (Eds.), *The genetics of neurological and psychiatric disorder* (pp. 115-120). New York, NY: Raven Press.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.).Toronto, ON: Pearson.

Thompson, B. (2004). Exploratory factor analysis decision sequence. *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* Washington, DC: American Psychology Association.

Thompson, J. K., & Cafri, G. (Eds.). (2007). *The muscular ideal: Psychological, social, and medical perspectives*. Washington, DC: American Psychological Association. doi:10.1037/11581-011

Thompson, M. A., & Gray, J. J. (1995). Development and validation of a new body-image assessment scale. *Journal of Personality Assessment*, *64*, 258-269. doi:10.1207/s1532 7752jpa6402_6

Thompson, J. K., Heinberg, M. A., & Tantleff-Dunn, S. (2002). *Exacting beauty: Theory, assessment, and treatment of body image disturbance*. Washington, DC: American Psychological Association.

Weinke, C. (1998). Negotiating the male body: Men, masculinity, and cultural ideals. *Journal of Men's Studies, 6*, 255-282. Retrieved from http://www.mensstudies.info/wordpress /?page_id=46

Willis, G., DeMaio, T., & Harris-Kojetin, B. (1999). Is the bandwagon headed to the methodological promise land? Evaluation of the validity of cognitive interviewing techniques. In M. Sirken, D. Hermann, S. Schechter, N. Schwartz, J. Tanur, & R. Tourangeau (Eds.), *Cognitive and survey research* (pp. 133-153). New York, NY: Wiley.

Yelland, C., & Tiggemann, M. (2003). Muscularity and the gay ideal: Body dissatisfaction and disordered eating in homosexual men. *Eating Behaviors*, *4*, 107-116. doi:10.1016/S1471-0153(03)00014-X

Appendix A

Paper 1: Recruitment Email

Dear Dr. *****,

I am inviting you to participate in a content validity study. You have been identified as an expert in the area of physical self-concept and/or measurement, and are being invited to participate in an on-line survey that will be part of my PhD dissertation. An information letter that provides a brief description of the study is included in the link below.

Your participation would include completing the on-line survey within the next 3 weeks. The survey will ask you about your expertise in this field and about the relevance and representativeness of newly developed items designed to assess muscularity, strength and body-fat as a part of physical self-concept. This survey should take approximately 15 minutes to complete.

If you decide to participate, simply click on the link to the information letter and the survey provided below. Your consent will be implied by your overt action of completing the survey.

http://fluidsurveys.com/surveys/christina/2r-expert-panel-survey/
If you have any questions about this email, you can contact me or my supervisor, Dr. Wendy Rodgers (wendy.rodgers@ualberta.ca), or the chair of the PER/ALES/NS REB at the University of Alberta listed in the information letter. I look forward to hearing from you.

Thanks,

Christina Loitz
Ph.D. Candidate
Faculty of Physical Education and Recreation,
University of Alberta,
Edmonton, Alberta, Canada
780-492-6899
christina.loitz@ualberta.ca

Appendix B

Study 1: Information Letter and Survey Package

Information Letter

Investigators: Christina Loitz, M.Sc., University of Alberta, 780 492 7424 Wendy Rodgers, Ph.D., Ph.D. Supervisor, University of Alberta, 780 492 2677 John Dunn, Ph.D., University of Alberta, 780 492 2831 Tanya Berry, Ph.D., University of Alberta, 780 492 3280

Title of project: Assessing the content relevance and representativeness of physical self-concept of muscle-tone, muscle-bulk, muscle-strength and body-fat items.

Dear Expert,

Thank you for considering participating in this study. I am interested in creating an instrument that assesses the muscle-tone, muscle-bulk, muscle-strength and body-fat as aspects of physical self-concept. These aspects of the physical self are not currently addressed in a single instrument. This is the topic of my doctoral dissertation.

The purpose of this study is to develop validity evidence pertaining to content relevance and content representativeness supporting the muscle-tone, muscle-bulk, muscle-strength and body-fat items. You have been selected as a prospective participant due to your expertise in the area of physical self-concept and/or measurement.

I am inviting you to participate in this study by completing an on-line survey that will take approximately 10- 15 minutes. The survey will include some demographic questions that will ask you about your gender, academic background, and whether you have publications or presentations related to physical self-concept. These questions will be followed by a set of questions asking about the relevance and representativeness of items reflecting muscle-tone, muscle-bulk, muscle-strength and body-fat.

The benefits of this research to you are minimal except that you will have an opportunity to contribute to the development of a new instrument and to the completion of my dissertation. The data from this study will be used as part of a dissertation project, for publications in academic journals and academic presentations. There are no expected risks associated with your participation in this study.

The information you provide in this study will be held in strict confidence. Only the researchers listed above will have access to your identifying information. As soon as the study is over, your name will be removed from all of the study materials and will be replaced with a numerical code. Your information will not be identifiable when you have completed your participation.

You are a volunteer in the study and we appreciate your participation. You are, of course, free to withdraw from the study without consequence at any time or to refuse to answer any question you do not wish to answer. If you have any questions about this study, you may contact any of the investigators listed above, or you may contact Dr. Kelvin Jones, 492 5910, of the Faculty of Physical Education and Recreation and Agricultural, Life and Environmental Sciences Research Ethics Board, who is not directly involved in this research.

Again, we thank you for your consideration.

Sincerely yours

Christina Loitz and the research team.

Appendix C

Study 1: Survey Package

PART 1: Demographic information:

Sex

☐ Male
☐ Female

What is your highest degree of education?

☐ Masters Degree
☐ Doctoral Degree
☐ Other Degree

What is the academic rank of your current position?

☐ Assistant professor
☐ Associate professor
☐ Professor
☐ Other
☐ Not applicable

Please describe your research area.

Have you published physical self-concept research in a peer reviewed setting?

○ Yes
○ No

Additional comments:

Have you presented physical self-concept research in a peer reviewed setting?

○ Yes
○ No

Additional comments:

Have you employed physical self-concept inventories in your research?

- Yes
- No

Additional comments:

Part 2: Content Relevance

Please indicate how well you feel each of the items reflects the domain specification it was written for. Judge the test item solely on the basis of the match between its content and the content defined by the domain specification that the test item was prepared to measure.

Section 1 - Muscle-tone Content Relevance

Muscle-tone

Operational Definition - The self-perception of the degree to which muscles look lean, firm and defined .

Item #1 - My muscles are toned.

Please rate the relevance of item#1 to muscle-tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#1:

Item #2 - My body is firm.

Please rate the relevance of item#2 to muscle-tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#2:

Item #3 - My muscles appear lean.

Please rate the relevance of item#3 to muscle-tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#3:

Item #4 - My muscles are well defined.

Please rate the relevance of item#4 to muscle-tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#4:

Item #5 - My body looks muscular.

Please rate the relevance of item#5 to muscle-tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#5:

Item #6 - I have a large amount of muscle mass on my body.

Please rate the relevance of item#6 to muscle-tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#6:

Item #7 - My body looks flabby.

Please rate the relevance of item#7 to muscle-tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#7:

Item #8 - My body is soft.

Please rate the relevance of item#8 to muscle-tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#8:

Section 2 - Muscle-bulk Content Relevance

Construct - Muscle-bulk

Operational Definition - The self-perception of the size and mass of muscles.

Item #9 - I have large muscles.

Please rate the relevance of item#9 to muscle-bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #9:

Item #10 - I have bulky muscles.

Please rate the relevance of item#10 to muscle-bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #10:

Item #11 - I have huge muscles.

Please rate the relevance of item#11 to muscle-bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #11:

Item #12 - I have small muscles.

Please rate the relevance of item#12 to muscle-bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #12:

Item #13 - I have scrawny muscles.

Please rate the relevance of item#13 to muscle-bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #13:

Item #14 - I have no muscles.

Please rate the relevance of item#14 to muscle-bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #14:

Item #15 - My muscles are slender.

Please rate the relevance of item#15 to muscle-bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #15:

Item #16 - I have bulging muscles.

Please rate the relevance of item#16 to muscle-bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #16:

Section 3 - Muscle-strength Content Relevance

Construct - Muscle-strength

Operational Definition - The self-perception of the presence or absence of strength and the ability to lift heavy objects.

Item #17 - I am physically strong.

Please rate the relevance of item#17 to muscle-strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #17:

Item #18 - I am physically weak.

Please rate the relevance of item#18 to muscle-strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #18:

Item #19 - I can lift heavy objects.

Please rate the relevance of item#19 to muscle-strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #19:

Item #20 - I cannot lift heavy objects.

Please rate the relevance of item#20 to muscle-strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #20:

Item #21 - I am physically powerful.

Please rate the relevance of item#21 to muscle-strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #21:

Item #22 - I am physically feeble.

Please rate the relevance of item#22 to muscle-strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #22:

Item #23 - My muscles are strong.

Please rate the relevance of item#23 to muscle-strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #23:

Item #24 - My muscles are weak.

Please rate the relevance of item#24 to muscle-strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #24:

Section 4 - Body-fat Content Relevance

Construct - Body-fat

Operational Definition - A self-perception relating to the amount of body-fat.

Item #25 - My body is fat.

Please rate the relevance of item#25 to body-fat.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #25:

Item #26 - My body is skinny.

Please rate the relevance of item#26 to body-fat.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item #26:

Item #27 - My body is large.

Please rate the relevance of item#27 to body-fat.

Poor match
Fair match
Good match
Very good match
Excellent match

Comments concerning item #27:

Item #28 - My body is thin.

Please rate the relevance of item#28 to body-fat.

Poor match
Fair match
Good match
Very good match
Excellent match

Comments concerning item #28:

Item #29 - I have a plus-size body.

Please rate the relevance of item#29 to body-fat.

Poor match
Fair match
Good match
Very good match
Excellent match

Comments concerning item #29:

Item #30 - I have a lean body.

Please rate the relevance of item#30 to body-fat.

Poor match
Fair match
Good match
Very good match
Excellent match

Comments concerning item #30:

Item #31 - I am chubby.

Please rate the relevance of item#31 to body-fat.

Poor match
Fair match
Good match
Very good match
Excellent match

Comments concerning item #31:

Item #32 - I have a plump body

Please rate the relevance of item#32 to body-fat.

Poor match
Fair match
Good match
Very good match
Excellent match

Comments concerning item #32:

PART 3: Content Representativeness

Content representativeness: Content representativeness refers to how well a group of items assesses all identified aspects of a construct.

Please rate the following items according to their representativeness to the constructs listed below.

Construct - Muscle-tone

Operational Definition - The self-perception of the degree to which muscles look lean, firm and defined.

Pool of items representing muscle-tone:

1. My muscles are toned.

2. My body is firm.

3. My muscles appear lean.

4. My muscles are well defined.

5. My body looks muscular.

6. My body lacks muscle definition.*

7. My body looks flabby.*

8. My body is soft.*

* reverse scoring

How well do you feel all the items included in the item pool represent the construct of muscle-tone?

- Poor Representativeness
- Fair Representativeness
- Good Representativeness
- Very Good Representativeness
- Excellent Representativeness

Are there any additional items that you feel should be included to represent the construct of muscle-tone?

- Yes
- No

Additional comments:

Are there any items you feel are redundant or could be eliminated from this set of items representing muscle-tone?

- Yes
- No

Additional comments:

Please rate the following items according to their representativeness to the constructs listed below.

Construct - Muscle-bulk

Operational Definition - The self-perception of the size and mass of muscles.

Pool of items representing muscle-bulk:

9. I have large muscles.

10. I have bulky muscles.

11. I have huge muscles.

12. I have small muscles.

13. I have scrawny muscles.*

14. I have no muscles.*

15. My muscles are slender.*

16. I have bulging muscles.

* reverse scoring

How well do you feel all the items included in the item pool represent the construct of muscle-bulk?

- Poor Representativeness
- Fair Representativeness
- Good Representativeness
- Very Good Representativeness
- Excellent Representativeness

Are there any additional items that you feel should be included to represent the construct of muscle-bulk?

- Yes
- No

Additional comments:

Are there any items you feel are redundant or could be eliminated from this set of items representing muscle-bulk?

- Yes
- No

Additional comments:

Please rate the following items according to their representativeness to the constructs listed below.

Construct - Muscle-strength

Operational Definition - The self-perception of the presence or absence of strength and the ability to lift heavy objects.

17. I am physically strong.

18. I am physically weak.*

19. I can lift heavy objects.

20. I cannot lift heavy objects.*

21. I am physically powerful.

22. I am physically feeble.*

23. My muscles are strong.

24. My muscles are weak.*

* reverse scoring

How well do you feel all the items included in the item pool represent the construct of muscle-strength?

- Poor Representativeness
- Fair Representativeness
- Good Representativeness
- Very Good Representativeness
- Excellent Representativeness

Are there any additional items that you feel should be included to represent the construct of muscle-strength?

- Yes
- No

Additional comments:

Are there any items you feel are redundant or could be eliminated from this set of items representing muscle-strength?

- Yes
- No

Additional comments:

Please rate the following items according to their representativeness to the constructs listed below.

Construct - Body-fat

Operational Definition - A self-perception relating to the amount of body-fat.

25. My body is fat.

26. My body is skinny.*

27. My body is large.

28. My body is thin.*

29. I have a plus-size body.

30. I have a lean body.*

31. I am chubby.

32. I have a plump body.

How well do you feel all the items included in the item pool represent the construct of body-fat?

- Poor Representativeness
- Fair Representativeness
- Good Representativeness
- Very Good Representativeness
- Excellent Representativeness

Are there any additional items that you feel should be included to represent the construct of body-fat?

- Yes
- No

Additional comments:

Are there any items you feel are redundant or could be eliminated from this set of items representing body-fat?

- Yes
- No

Additional comments:

Thank you for completing the 2R Expert Panel Survey. I appreciate the time that you have spent completing this survey. If you have any comments on the survey or the process please comment below or you can send me an email at christina.loitz@ualberta.ca

Appendix D

Study 2: Summary of the Qualitative and Quantitative of the Sixteen Judges' Ratings of Content Relevance of the General Physical Self-Concept Scale

Construct	Item	Comments
Muscle-tone	1 My muscles are toned.	- maybe too general, it doesn't capture a specific facet of muscle-tone -(see comment under #3) I'm contrasting this with Item #1 and wondering if you are interested in whether or not muscles "appear" toned or "are" toned. I think there is a potential difference. Based on your operational definition, the items that tap in to "appearing" toned are closer to that definition than the items that say "is" or "are" toned. Just something to consider. Whether the items need to change or the operational definition needs to be more inclusive, I'm not sure. (Also, sorry if some of these comments are grammatically awkward or there are spelling mistakes. I can't actually see the type that I put in because the box is so small and the text just runs through it...)
	2 My body is firm.	- I wonder if a firm body necessarily means toned muscles. Can someone have a firm body without a lot of muscle-tone? I'd say yes. - The whole body? Just relevant to muscles? - There is no reference to tone -- I could be firm but not defined and thus Excellent or poor -- and no form of comparison -Not necessarily related to muscles specifically - perception of what firm means can vary - may be too vague or interpreted as 'not fat' - Firmness seem as function of leanness (low fat) as well as muscle-tone
	3 My muscles appear lean.	- I'm contrasting this with Item #1 and wondering if you are interested in whether or not muscles "appear" toned or "are" toned. I think there is a potential difference. Based on your operational definition, the items that tap in to "appearing" toned are closer to that definition than the items that say "is" or "are" toned. Just something to consider. Whether the items need to change or the operational definition needs to be more inclusive, I'm not sure. (also, sorry if some of these comments are grammatically awkward or there are spelling mistakes. I can't actually see the type that I put in because the box is so small and the text just runs through it...) - Appear to me (or others) - I am not sure someone would describe 'muscles' as being lean - one's body appears lean - To me lean is about lack of fat so how can muscles look lean? this could be confusing - Lean muscles appear to refer to muscles that are small in size (many would think of a ballerina). One would need to add 'lean' with tone to describe this particular type of body appearance
	4 My muscles are well defined.	- Related to muscularity rather than 'tone' - This is a little more generic - (same as #3) I'm contrasting this with Item #1 and wondering if you are interested in whether or not muscles "appear" toned or "are" toned. I think there is a potential difference. Based on your operational definition, the items that tap in to "appearing" toned are closer to that definition than the items that say "is" or "are" toned. Just something to consider. Whether the items need to change or the operational definition needs to be more inclusive, I'm not sure. (Also, sorry if some of these comments are grammatically awkward or there are spelling mistakes. I can't actually see the type that I put in because the box is so small and the text just runs through it...)

(continued)

Construct	Item	Comments
	5 My body looks muscular.	-This item does not differentiate the appearance of having lean, toned muscles from having big bulky muscles - This is one item that will need some specific thought because the multiple meanings of muscularity are at stake here. Just having "muscular" could mean tone or bulk depending on whose reading it and how they interpret it. I.e., this item will have some domain overrepresentation to it. - Yes, but not toned - May relate more to size than tone - May not be specific to 'tone', too vague maybe - Related to muscularity rather than 'tone'
	6 I have a large amount of muscle mass on my body.	- This items is clearly about amount or size of muscles rather than a lean, toned appearance - My musculature is greater than others of my gender and age? - Muscle mass and tone are different things and this item is focused on a "large amount of muscle mass" so... - This has more to do with size then definition/tone - It's about bulk - Large amount - not sure about that phrase - Does not necessarily imply lean or toned muscles - This seems more like bulk
	7 My body looks flabby.	- This item does not address muscle-tone. Also, it would be a poor negatively scored item for muscle-tone because it is possible to believe you do not appear muscular without thinking you look fat for flabby. - Be careful you are not just asking the opposite of my body is toned. What you are moving towards are bloated specifics (Catell, 19**) - I think that this and item #8 could work, but I'm not sure the flabbiness and softness is necessarily equivalent to muscularity. But it might be worth keeping to see how they hang together with the other items. In principle though I support the idea of trying to have different kinds of items (rather than basically just saying the same thing over and over, which would be my concern with items 1-5 (do you really need all 5 of them - or will participants get frustrated that you just keep asking them the same question over and over). -Reverse scored I assume. Reverse scoring tends to enter method effects into measurement that may not be entirely desirable. So, even though I've said "good match", that rating comes with a caveat. - Obviously more related to fat but you cannot look toned with lots of fat so may work - Seems to be more relevant to body-fat
	8 My body is soft.	- This item does not address muscle-tone. Also, it would be a poor negatively scored item for muscle-tone because it is possible to believe you do not appear muscular without thinking you look soft. For example, you could think you look very skinny. - Muscles not body? - Reverse scored I assume. Reverse scoring tends to enter method effects into measurement that may not be entirely desirable. So, even though I've said "good match", that rating comes with a caveat. - I think many may relate this to skin texture - I would expect a gender bias to this question
Muscle-bulk	9 I have large muscles.	- My comments would be similar to the other section - I think the operational definition needs to be more clear as to whether it is "appear" to have muscle-bulk or "has" muscle-bulk. Then the items need to reflect that choice. And that choice would influence the match I see in the items you've presented. - Tone was defined in terms of the looks of the body - bulk does not refer to how the body looks?

(continued)

Construct	Item	Comments
	10 I have bulky muscles.	- While this item uses the same working as the title of the construct, the matched operational definition seems problematic to me. Specifically, perceiving that one has large muscles or substantial muscle mass seems to connote a positive perception for someone who values large muscles. However, to say those muscles are, or look bulky has a more negative connotation. I believe this distinction is subtle, which is why I rated it a good match rather than fair or poor. - Bulky - might create some challenges in interpretation - (same as above) My comments would be similar to the other section - I think the operational definition needs to be more clear as to whether it is "appear" to have muscle-bulk or "has" muscle-bulk. Then the items need to reflect that choice. And that choice would influence the match I see in the items you've presented.
	11 I have huge muscles.	- (same as two above) All you are doing is asking the same question but in magnitude each time BLOATED SPECIFICS - May work to define extremes
	12 I have small muscles.	- Good as a reverse scored item - Opposite of #9 - I'm getting a bit concerned that all of the items are asking the same thing...I think you'll get a high alpha - but are the multiple items really needed doing it this way. This is a very difficult task in scale development, but I encourage you to really think about it - otherwise my fear is that you'll end up with a 20-item, 5 subscale scale with the 5 items in each subscale essentially just being repetitive. It might end with good measurement properties, but I think would add unnecesary burdon to your participants - and I'm not sure you'd really be pleased with the end result of the scale (and your reviewers when you submit for publication might pick up on this and just say all you've done is repeated the same question over and over). Sorry to ramble, but based on my experience with scales I strongly recommend at least thinking about this issue as you're putting the scale together. You're putting a lot of work into this scale development, and you don't want to end up with something that someone could just as easily assess using 4 items that basically just ask a question like "I have toned muscles" for each domain.
	13 I have scrawny muscles.	- See my earlier comments on reverse scoring of items. - 'scrawny' north American term - OK as a reverse scored item. Again I think the negative evaluate tone of this wording compared to items that focus on a description of the size or mass of muscles may be problematic.
	14 I have no muscles.	- This item does not seem adequately distinct from muscle-tone. - Impossible to have no muscles. You would be just a bag of bones without them - Impossible? - See my earlier comments on reverse scoring of items. - May work as an extreme - Not possible - It isn't realistic to say "no" muscles - Everybody has "muscle'---just the extent to which it is perceived to self (and or others) as having 'size'

(continued)

Construct	Item	Comments
	15 My muscles are slender.	- Very good as a reverse scored item - Not a big fan of "slender", but this might just be a personal preference. - See my earlier comments on reverse scoring of items. As well, for some people this might be confounded with muscle-tone. - Seems more relevant to muscle-tone and also gender bias - Good match for definition, but 'slender' doesn't seem like an appropriate term - Is this like toned? - Muscles aren't often thought of as being slender
	16 I have bulging muscles.	- Sounds like strange choice of wording - bulk seems better - I think muscles could be big, but not necessarily bulging. To me, bulging includes the shape moreso than just bulk - heading it a bit towards toneness. - This does not seem entirely distinct from muscle-tone. It also may be problematic if bulging is interpreted as connoting a negative evaluation.
Muscle-strength	17 I am physically strong.	- Muscle-strength is quite different from 'tone' or 'bulk' that refer to the looks of the body. Strength refers person ability to do things (e.g., lift). It is also an objectively measurable, unlike tone. Therefore, it does not work as well as a 'self-perception' item.
	18 I am physically weak.	- Very good as a reverse scored item - See my earlier comments on problems with reverse scored items. - Opposite of 17 - Negative items are always a challenge
	19 I can lift heavy objects.	- "heavy" is open to individual interpretation - I am strong enough to lift ... - Give example of what heavy would mean - I think people might think more about injury in this case, not just pure strength. - Can be skinny but strong
	20 I cannot lift heavy objects.	- Opposite of 19. no new info will be gained here - See my earlier comments on problems with reverse scored items. - Can be limited by back pain etc - (see comment from #19) give example of what heavy would mean - (see comment from #19) "heavy" is open to individual interpretation
	21 I am physically powerful.	- May be more related to size - Power and strength are often used colloquially to mean the same thing--but some people are aware that strength is the ability to generate force while power is the ability to generate force quickly. This may result in a confound. - The knowledgeable consumer will differentiate between power and strength. Even myself, I'd say I have pretty good muscle-strength, but my self-perception of power would be much lower. - Repeat
	22 I am physically feeble.	- Reverse scored - Feeble can have different meanings - I think this would include more than just strength. - See my earlier comments on problems with reverse scored items.

(continued)

Construct	Item	Comments
	23 My muscles are strong.	- Repeat
	24 My muscles are weak.	- Reverse scored - No new info in asking this - See my earlier comments on problems with reverse scored items.
Body-fat	25 My body is fat.	- This may be interpreted more as shape than actual body-fat - Similar to strength, body-fat is objectively measurable and thus, is not dependent on self-perception only. It is different to measure body-fat (%) that to ask if individuals perception themselves fat. You might need to change this item 'fatness' as often the body-fat% has not much to with the perception of fatness
	26 My body is skinny.	- See my earlier comments on problems with reverse scored items. - Opposite of 25 - Reverse scored
	27 My body is large.	- Does not distinguish large muscles from large and fat. - I think this will get very confused with muscle-bulk. - Could be mistaken for height - Could be large without fat - Large doesn't discriminate between largeness as a consequence of adiposity and and largeness as a consequence of muscularity - Is large shape or muscularity? - Confounded with bulk - Confounded with large muscular bodies - Large is not necessarily fat - Large could be interpreted as having muscle mass (or being tall) as well
	28 My body is thin.	- Reverse scored - Same as skiiny - See my earlier comments on problems with reverse scored items.
	29 I have a plus-sized body.	- In this case it may be OK, but body size is not always interpreted as body-fatness. - Good if this is for females only - Plus size is a north amercian term- - Not as phrase I am familiar with - Gender bias - a term not often used by males - Good match - but, men might not know how to respond to this - Perhaps gender biased - (see comment for #27)I think this will get very confused with muscle-bulk.
	30 I have a lean body.	- May be some overlap with muscle-tone - This is actually an interesting one. Maybe the leanness and body-fat scales are one and the same...(at least from a measurement perspective) - Leanness can sometimes be confounded with tone. - More relevant to body tone

(continued)

Construct	Item	Comments
	31 I am chubby.	No comments
	32 I have a plump body.	No comments

Appendix E

Study 2: Survey Package

Title of project: Assessing the content relevance and representativeness of physical self-concept of muscle tone, muscle bulk, muscle strength and body fat items.

PART 1: Demographic information:

Sex

☐ Male
☐ Female

What is your highest degree of education?

☐ Certificate
☐ Diploma
☐ Undergraduate Degree
☐ Masters Degree
☐ Doctoral Degree
☐ Other Degree

What are your current fitness credentials?

☐ AFLCA
☐ CSEP CPT
☐ CSEP CEP
☐ Other: ____

Please describe the work do you do in the area of health and fitness?

____
____

Please describe your fitness clients?

____
____

Part 2: Content Relevance

Please indicate how well you feel each of the items reflects the term it was written for. Judge the question solely on the basis of the match between its content and the content defined by the operational definition that the question it was prepared to measure.

Section 1 - Muscle Tone Content Relevance

Muscle Tone

Operational Definition - The self-perception of the degree to which muscles appear lean, firm and defined.

Comments:____

Item #1 - My muscles appear toned.

Please rate the relevance of item#1 to muscle tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#1:

____
____

Item #2 - My muscles appear firm.

Please rate the relevance of item#2 to muscle tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#2:

____
____

Item #3 - My muscles appear lean.

Please rate the relevance of item#3 to muscle tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#3:

Item #4 - My muscles appear well defined.

Please rate the relevance of item#4 to muscle tone.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#4:

<u>Section 2 - Muscle Bulk Content Relevance</u>

Construct - Muscle Bulk

Operational Definition - The self-perception of the degree to which <u>muscles appear large and bulky in size.</u>

Comments:____

Item #5 - I appear to have large muscles.

Please rate the relevance of item#5 to muscle bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#5:

____

Item #6 – My muscles appear bulky.

Please rate the relevance of item#6 to muscle bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#6:

____

Item #7 - My muscles appear to be huge.

Please rate the relevance of item#7 to muscle bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#7:

____

Item #8 – My muscles appear small.

Please rate the relevance of item#8 to muscle bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#8:

____
____

Item #9 - I appear to have bulging muscles.

Please rate the relevance of item#9 to muscle bulk.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#9:

____
____

<u>Section 3 - Muscle Strength Content Relevance</u>

Construct - Muscle Strength

Operational Definition - The self-perception of the <u>presence or absence of muscle strength</u> and the ability to lift heavy objects.

Comments:____

Item #10 - I am physically strong.

Please rate the relevance of item#10 to muscle strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#10:

____
____

Item #11 - I am physically weak.

Please rate the relevance of item#11 to muscle strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#11:

____
____

Item #12 - I am strong enough to lift heavy objects.

Please rate the relevance of item#12 to muscle strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#12:

____
____

Item #13 - I am physically feeble.

Please rate the relevance of item#13 to muscle strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#13:

____
____

Item #14 - My muscles are strong.

Please rate the relevance of item#14 to muscle strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#14:

____
____

Item #15 - My muscles are weak.

Please rate the relevance of item#15 to muscle strength.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#15:

____
____

Section 4 - Body Fat Content Relevance

Construct - Body Fat

Operational Definition - The self-perception of the degree of fat that appears on the body.

Comments:____

Item #16 - My body appears fat.

Please rate the relevance of item#16 to body fat.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#16:

____

Item #17 - My body appears skinny.

Please rate the relevance of item#17 to body fat.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#17:

____

Item #18 - My body appears thin.

Please rate the relevance of item#18 to body fat.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#18:

____

Item #19 - My body appears lean.

Please rate the relevance of item#19 to body fat.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#19:

____
____

Item #20 - I appear chubby.

Please rate the relevance of item#20 to body fat.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#20:

____
____

Item #21 - I appear to have a plump body

Please rate the relevance of item#21 to body fat.

- Poor match
- Fair match
- Good match
- Very good match
- Excellent match

Comments concerning item#21:

____
____

PART 3: Content Representativeness

Content representativeness: Content representativeness refers to how well a group of items assesses all identified aspects of a construct.

Please rate the following items according to their representativeness to the constructs listed below.

Construct - Muscle Tone

Operational Definition - The self-perception of the degree to which muscles appear lean, firm and defined.

Pool of items representing muscle tone:

Item #1 - My muscles appear toned.

Item #2 - My muscles appear firm.

Item #3 - My muscles appear lean.

Item #4 - My muscles appear well defined.

How well do you feel all the items included in the item pool represent the construct of muscle tone?

O Poor representation
O Fair Representation
O Good Representation
O Very Good Representation
O Excellent Representation

Comment:____

What words do you or your clients use to talk about muscle tone:____

Construct - Muscle Bulk

Operational Definition - The self-perception of the degree to which muscles appear large and bulky in size.

Item #5 - I appear to have large muscles.

Item #6 – My muscles appear bulky.

Item #7 - My muscles appear to be huge.

Item #8 – My muscles appear small. *

Item #9 - I appear to have bulging muscles.

* reverse scored

How well do you feel all the items included in the item pool represent the construct of muscle bulk?

O Poor representation

O Fair Representation

O Good Representation

O Very Good Representation

O Excellent Representation

Comment:____

What words do you or your clients use to talk about muscle bulk:____

Construct - Muscle Strength

Operational Definition - The self-perception of the presence or absence of muscle strength and the ability to lift heavy objects.

Item #10 - I am physically strong.

Item #11 - I am physically weak. *

Item #12 - I am strong enough to lift heavy objects.

Item #13 - I am physically feeble. *

Item #14 - My muscles are strong.

Item #15 - My muscles are weak.

*** reverse scoring**

How well do you feel all the items included in the item pool represent the construct of muscle strength?

O Poor representation

O Fair Representation

O Good Representation

O Very Good Representation

O Excellent Representation

Comment:____
____

What words do you or your clients use to talk about muscle strength:____
____

Construct - Body Fat

Operational Definition - The self-perception of the degree of fat that appears on the body.

Item #16 - My body appears fat.

Item #17 - My body appears skinny.

Item #18 - My body appears thin.

Item #19 - I body appears lean.

Item #20 - I appear chubby.

Item #21 - I appear to have a plump body

How well do you feel all the items included in the item pool represent the construct of body fat?

O Poor representation

O Fair Representation

O Good Representation

O Very Good Representation

O Excellent Representation

Comment:____

____

What words do you or your clients use to talk about body fat:

____

____

Appendix F

Study 2: Email Invitation

Re: Assessing the content relevance and representativeness of physical self-concept of muscle-tone, muscle-bulk, muscle-strength and body-fat items.

Dear_________________,

I am inviting you to participate in a study that involves the development of a new survey to assess perceptions of muscularity, body-fatness and strength. You have been identified as an expert in the area of physical fitness, and are being invited to participate in a study that will be part of my PhD dissertation. An information letter that provides a brief description of the study is attached to this email.

Your participation would include completing a short pen and paper survey within the next 3 weeks. I will then ask you a few questions about your responses to the survey. The survey will ask you about your fitness expertise and your thoughts about a new survey that has been designed to measure self-perceptions of muscularity, strength and body-fat. This session should take approximately 30 minutes to complete.

If you decide to participate, simply reply to this email and we will book a time to meet.

If you have any questions about this email, you can contact me or my supervisor, Dr. Wendy Rodgers (wendy.rodgers@ualberta.ca), or the chair of the PER/ALES/NS REB at the University of Alberta listed in the information letter. I look forward to hearing from you.

Thanks,
Christina Loitz
Ph.D. Candidate
Faculty of Physical Education and Recreation,
University of Alberta,
Edmonton, Alberta, Canada
780-492-6899
christina.loitz@ualberta.ca

Appendix G

Study 2: Information Letter

Investigators: Christina Loitz, M.Sc., University of Alberta, 780 492 7424 ,
Wendy Rodgers, Ph.D., Ph.D. Supervisor, University of Alberta, 780 492 2677
John Dunn, Ph.D., University of Alberta, 780 492 2831, and
Tanya Berry, Ph.D., University of Alberta, 780 492 3280

Title of project: Assessing the content relevance and representativeness of physical self-concept of muscle-tone, muscle-bulk, muscle-strength and body-fat items.

Dear Fitness Expert,

Thank you for considering participating in this study. I am interested in creating a questionnaire that assesses self-perceptions of muscle-tone, muscle-bulk, muscle-strength and body-fat. These aspects of the physical self are not currently addressed in a single instrument. The purpose of this study is to get feedback from fitness experts (personal trainers or fitness instructors) about questions concerning peoples' perceptions of their muscle-tone, muscle-bulk, muscle-strength and body-fat. You have been selected as a prospective participant due to your expertise and experience as a personal trainer and/or fitness instructor.

I am inviting you to participate in this study by completing a short oral survey. The survey will include some questions about your gender, type of fitness expertise and experiences as a fitness professional. This will be followed by a set of questions asking about the degree of match between the questions that are proposed to be included on a new survey reflecting muscle-tone, muscle-bulk, muscle-strength and body-fat. You will be asked some follow-up questions about how you responded to these questions. This will be recorded using an audio digital recorder.

The benefits of this research to you are minimal except that you will have an opportunity to contribute to the development of a new questionnaire and to the completion of my dissertation. The data from this study will be used as part of a dissertation project, for publications in academic journals and academic presentations. There are no expected risks associated with your participation in this study.

The information you provide in this study will be held in strict confidence. Only the researchers listed above will have access to your identifying information. As soon as the study is over, your name will be removed from all of the study materials and will be replaced with a numerical code. Your information will not be identifiable when you have completed your participation.

You are a volunteer in the study and we appreciate your participation. You are, of course, free to withdraw from the study without consequence at any time or to refuse to answer any question you do not wish to answer. If you have any questions about this study, you may contact any of the investigators listed above, or you may contact Dr. Kelvin Jones, 492 5910, of the Faculty of Physical Education and Recreation and Agricultural, Life and Environmental Sciences Research Ethics Board, who is not directly involved in this research. Again, we thank you for your consideration.

Sincerely yours,

Christina Loitz and the research team

Appendix H

Study 3: Survey Package

PART 1 –Demographic Questions

1. What is your gender? Please circle one of the following. Male / Female
2. What is your current age? ______ years old
3. What is your current height? _____________
4. What is your current weight? _____________

Please answer the next questions according to what you ***did*** over the last 4 weeks.

1. Considering a **7-day period** (a week), how many times on the average did you do the following kinds of exercise for a **15 minutes** bout during your **free time** (write in each square the appropriate number).

TIMES PER WEEK

a) **Strenuous Exercise *(Heart beats rapidly)***
(i.e. running, jogging, hockey, football, soccer, squash, basketball, cross country skiing, judo, roller skating, vigorous swimming, vigorous long distance bicycling, strength training)

b) **Moderate Exercise *(Not exhausting)***
(i.e. fast walking, baseball, tennis, easy bicycling, volleyball, badminton, easy swimming, alpine skiing, popular and folk dancing)

c) **Mild Exercise *(Minimal effort)***
(i.e. yoga, archery, fishing from the riverbank, bowling, horseshoes, golf, snowmobiling, easy walking)

PART 2 – Overall Physical Self-Concept

Instructions: The purpose of this questionnaire is to identify how you view your body. Please help us to more fully understand how people feel about their bodies by indicating the extent to which you agree or disagree with the statements below. Please circle one response option (please do not circle a spot between the response options). ***There are no right or wrong answers*** so please don't spend too much time on any one statement; simply choose the answer that best describes how you view each statement.

Please mark on the scale below, which ranges from 1 "strongly disagree" to 5 "strongly agree", the extent that you agree or disagree with the following statements. Thanks.

1. My **muscles appear toned.**

1	2	3	4	5
Strongly disagree				Strongly agree

2. **I appear to have large muscles**.

1	2	3	4	5
Strongly disagree				Strongly agree

3. I am physically strong.

1	2	3	4	5
Strongly disagree				Strongly agree

4. My body appears fat.

1	2	3	4	5
Strongly disagree				Strongly agree

5. **My muscles appear firm.**

1	2	3	4	5
Strongly disagree				Strongly agree

6. **My muscles appear to be huge.**

1	2	3	4	5
				Strongly agree
Strongly disagree				

7. I am physically weak.

1	2	3	4	5
Strongly disagree				Strongly agree

8. My muscles appear bulky.

1	2	3	4	5
Strongly disagree				Strongly agree

9. My muscles are strong.

1	2	3	4	5
Strongly disagree				Strongly agree

10. I appear chubby.

1	2	3	4	5
Strongly disagree				Strongly agree

11. My muscles appear well-defined.

1	2	3	4	5
Strongly disagree				Strongly agree

12. My muscles are weak.

1	2	3	4	5
Strongly disagree				Strongly agree

13. My body jiggles.

1	2	3	4	5
Strongly disagree				Strongly agree

14. My muscles are ripped.

1	2	3	4	5
Strongly disagree				Strongly agree

15. **My muscles appear big**.

1	2	3	4	5
Strongly disagree				Strongly agree

16. **My body is flabby**.

1	2	3	4	5
Strongly disagree				Strongly agree

17. **My muscles are well sculpted**.

1	2	3	4	5
Strongly disagree				Strongly agree

Thank you for participating!

Appendix I

Seventeen General Physical Self-Concept Scale Items assessed in Study 3 and 4

General Physical Self-Concept Items

Item	Construct
1. My muscles appear toned.	Muscle-tone
2. I appear to have large muscles.	Muscle-bulk
3. I am physically strong.	Muscle-strength
4. My body appears fat.	Body-fat
5. My muscles appear firm.	Muscle-tone
6. My muscles appear to be huge.	Muscle-bulk
7. I am physically weak. *	Muscle-strength
8. My muscles appear bulky.	Muscle-bulk
9. My muscles are strong.	Muscle-strength
10. I appear chubby.	Body-fat
11. My muscles appear well-defined.	Muscle-tone
12. My muscles are weak.*	Muscle-strength
13. My body jiggles.	Body-fat
14. My muscles are ripped.	Muscle-tone
15. My muscles appear big.	Muscle-bulk
16. My body is flabby.	Body-fat
17. My muscles are well sculpted.	Muscle-tone

Note: * = item to be reverse scored. Total of 17 items included in Studies 3 and 4. The item numbers are the same in Study 3 and Study 4.

Appendix J

Study 4: Survey Package Time-1

PART 1 –Demographic Questions

1. Name:

2. Date:

3. What is your gender? Please circle one of the following.

   Male / Female

4. What is your current age? ______ years old

5. What is your current height? _____________

6. What is your current weight? _____________

7. What faculty are you in? ___________________________________

8. What is your major? ____

9. What ethnic or cultural group(s) do your ancestors belong to?

   ____

10. What ethnic or cultural group(s) do you belong to?

    ____

Please answer the next questions according to what you ***did*** over the last 4 weeks.

11. Considering a **7-day period** (a week), how many times on the average did you do the following kinds of exercise for a **15 minutes** bout during your **free time** (write in each square the appropriate number).

TIMES PER WEEK

a) **Strenuous Exercise *(Heart beats rapidly)***
(i.e. running, jogging, hockey, football, soccer, squash,
basketball, cross country skiing, judo, roller skating, vigorous swimming, vigorous long distance bicycling, strength training) ☐

b) **Moderate Exercise *(Not exhausting)***
(i.e. fast walking, baseball, tennis, easy bicycling, volleyball, badminton, easy swimming, alpine skiing, popular and folk dancing) ☐

c) **Mild Exercise *(Minimal effort)***
(i.e. yoga, archery, fishing from the riverbank, bowling, horseshoes, golf, snowmobiling, easy walking) ☐

Please circle the most accurate response:

12. In general, How many days per week do you do strength training activity?

0 days	1 day	2 days	3 days	4 days	5 days	6 days	7 days

13. In general, how many days per week do you do cardiovascular activity?

0 days	1 day	2 days	3 days	4 days	5 days	6 days	7 days

14. In general, how many days per week do you do flexibility activity?

0 days	1 day	2 days	3 days	4 days	5 days	6 days	7 days

PART 2 – OVERALL PHYSICAL SELF-CONCEPT

Instructions: The purpose of this questionnaire is to identify how you view your body. Please help us to more fully understand how people feel about their bodies by indicating the extent to which you agree or disagree with the statements below. Please circle one response option (please do not circle a spot between the response options). ***There are no right or wrong answers*** so please don't spend too much time on any one statement; simply choose the answer that best describes how you view each statement.

Please mark on the scale below, which ranges from 1 "strongly disagree" to 5 "strongly agree", the extent that you agree or disagree with the following statements. Thanks.

1. My muscles appear toned.

1	2	3	4	5
Strongly disagree				Strongly agree

2. I appear to have large muscles.

1	2	3	4	5
Strongly disagree				Strongly agree

3. I am physically strong.

1	2	3	4	5
Strongly disagree				Strongly agree

4. My body appears fat.

1	2	3	4	5
Strongly disagree				Strongly agree

5. My muscles appear firm.

1	2	3	4	5
Strongly disagree				Strongly agree

6. My muscles appear to be huge.

1	2	3	4	5
Strongly disagree				Strongly agree

7. I am physically weak.

1 Strongly disagree	2	3	4	5 Strongly agree

8. My muscles appear bulky.

1 Strongly disagree	2	3	4	5 Strongly agree

9. My muscles are strong.

1 Strongly disagree	2	3	4	5 Strongly agree

10. I appear chubby.

1 Strongly disagree	2	3	4	5 Strongly agree

11. My muscles appear well-defined.

1 Strongly disagree	2	3	4	5 Strongly agree

12. My muscles are weak.

1 Strongly disagree	2	3	4	5 Strongly agree

13. My body jiggles.

1 Strongly disagree	2	3	4	5 Strongly agree

14. My muscles are ripped.

1 Strongly disagree	2	3	4	5 Strongly agree

15. My muscles appear big.

1 Strongly disagree	2	3	4	5 Strongly agree

16. My body is flabby.

1 Strongly disagree	2	3	4	5 Strongly agree

17. My muscles are well sculpted.

1 Strongly disagree	2	3	4	5 Strongly agree

PART 3: Use the scale provided to rate the extent to which each item applies to you.

1. I consider myself an exerciser.

Strongly disagree — Strongly agree

1 2 3 4 5 6 7

2. When I describe myself to others, I usually include my involvement in exercise.

Strongly disagree — Strongly agree

1 2 3 4 5 6 7

3. I have numerous goals related to exercising.

Strongly disagree — Strongly agree

1 2 3 4 5 6 7

4. Exercise is a central factor to my self concept.

Strongly disagree							Strongly agree
	1	2	3	4	5	6	7

5. I need to exercise to feel good about myself.

Strongly disagree							Strongly agree
	1	2	3	4	5	6	7

6. Others see me as someone who exercises regularly.

Strongly disagree							Strongly agree
	1	2	3	4	5	6	7

7. For me, being an exerciser means more than just exercising.

Strongly disagree							Strongly agree
	1	2	3	4	5	6	7

8. I would feel a real loss if I were forced to give up exercising.

Strongly disagree							Strongly agree
	1	2	3	4	5	6	7

9. Exercising is something I think about often.

Strongly disagree							Strongly agree
	1	2	3	4	5	6	7

PART 4: Please read each question carefully and then, for each one, circle the number that best applies to you.

1. I wish that I were more muscular.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

2. I lift weights to build up muscle.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

3. I use protein or energy supplements.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

4. I drink weight-gain or protein shakes.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

5. I try to consume as many calories as I can in one day.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

6. I feel guilty if I miss a weight training session.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

7. I think I would feel more confident if I had more muscle mass.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

8. Other people think I work out with weights often.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

9. I think that I would look better if I gained 10 pounds in bulk.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

10. I think about taking anabolic steroids.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

11. I think that I would feel stronger if I gained a little more muscle mass.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

12. I think that my weight training schedule interferes with other aspects of my life.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

13. I think that my arms are not muscular enough.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

14. I think that my chest is not muscular enough.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

15. I think that my legs are not muscular enough.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

Part 5: Please read each question carefully and then, for each one, circle the number that best applies to you.

1. I eat sweets and carbohydrates without feeling nervous.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

2. I think about dieting.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

3. I feel extremely guilty after overeating.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

4. I am terrified of gaining weight.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

5. I exaggerate or magnify the importance of weight.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

6. I am preoccupied with the desire to be thinner.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

7. If I gain a pound, I worry that I will keep gaining.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

8. I think that my stomach is too big.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

9. I think my thighs are too large.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

10. I think that my stomach is just the right size.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

11. I feel satisfied with the shape of my body.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

12. I like the shape of my buttocks.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

13. I think my hips are too big.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

14. I think that my thighs are just the right size.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

15. I think my buttocks are too large.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

16. I think that my hips are just the right size.

Always	Very often	Often	Sometimes	Rarely	Never
1	2	3	4	5	6

PART 5: Please read the question carefully and then circle the number that best applies to you.

1) I intend to do strength training exercises over the next week.

1	2	3	4	5	6	7
Strongly disagree						Strongly agree

2) I intend to do cardiovascular exercises over the next week.

1	2	3	4	5	6	7
Strongly disagree						Strongly agree

3) I intend to do flexibility exercises over the next week.

1	2	3	4	5	6	7
Strongly disagree						Strongly agree

Thank you for participating.

Appendix K

Final General Physical Self-Concept Scale Instrument from Dissertation.

PART 1 GPSCS – MB, MS and BF

Instructions: The purpose of this questionnaire is to identify how you view your body. Please help us to more fully understand how people feel about their bodies by indicating the extent to which you agree or disagree with the statements below. Please circle one response option (please do not circle a spot between the response options). ***There are no right or wrong answers*** so please don't spend too much time on any one statement; simply choose the answer that best describes how you view each statement.

Please mark on the scale below, which ranges from 1 "strongly disagree" to 5 "strongly agree", the extent that you agree or disagree with the following statements. Thanks.

1. I appear to have large muscles.

1	2	3	4	5
Strongly disagree				Strongly agree

2. I am physically strong.

1	2	3	4	5
Strongly disagree				Strongly agree

3. My body appears fat.

1	2	3	4	5
Strongly disagree				Strongly agree

4. My muscles appear to be huge.

1	2	3	4	5
Strongly disagree				Strongly agree

5. I am physically weak.

1	2	3	4	5
Strongly disagree				Strongly agree

6. My muscles appear bulky.

1	2	3	4	5
Strongly disagree				Strongly agree

7. My muscles are strong.

1	2	3	4	5
Strongly disagree				Strongly agree

8. I appear chubby.

1	2	3	4	5
Strongly disagree				Strongly agree

9. My muscles are weak.

1	2	3	4	5
Strongly disagree				Strongly agree

10. My body jiggles.

1	2	3	4	5
Strongly disagree				Strongly agree

11. My muscles are ripped.

1	2	3	4	5
Strongly disagree				Strongly agree

12. My muscles appear big.

1	2	3	4	5
Strongly disagree				Strongly agree

13. My body is flabby.

1	2	3	4	5
Strongly disagree				Strongly agree

Part 2: GPSCS – MT

Instructions: The purpose of this questionnaire is to identify how you view your body. Please help us to more fully understand how people feel about their bodies by indicating the extent to which you agree or disagree with the statements below. Please circle one response option (please do not circle a spot between the response options). ***There are no right or wrong answers*** so please don't spend too much time on any one statement; simply choose the answer that best describes how you view each statement.

Please mark on the scale below, which ranges from 1 "strongly disagree" to 5 "strongly agree", the extent that you agree or disagree with the following statements. Thanks.

1. My muscles appear toned.

1	2	3	4	5
Strongly disagree				Strongly agree

2. My muscles appear firm.

1	2	3	4	5
Strongly disagree				Strongly agree

3. My muscles appear well-defined.

1	2	3	4	5
Strongly disagree				Strongly agree

4. My muscles are well sculpted.

1	2	3	4	5
Strongly disagree				Strongly agree