

9-1-2011

The Effects of the Type of Skill Test, Choice, and Gender on the Situational Motivation of Physical Education Students

Tyler G. Johnson
Boise State University

Keven A. Prusak
Brigham Young University

Todd Pennington
Brigham Young University

Carol Wilkinson
Brigham Young University

The Effects of the Type of Skill Test, Choice, and Gender on the Situational Motivation of Physical Education Students

Tyler G. Johnson¹, Keven A. Prusak², Todd Pennington²,
and Carol Wilkinson²

¹Boise State University, ²Brigham Young University

The purpose of this study was to examine the effects of (a) skill test type, (b) choices, and (c) gender on the situational motivation profiles of adolescents during skill testing in physical education. Participants were 507 students (53% male) aged 12–16 years ($M = 13.87$; $SD = 0.94$) attending a suburban junior high school in a western state in the U.S. All participants experienced either a norm-referenced, summative or a criterion-referenced, formative skill test with or without choices. The Situational Intrinsic Motivation Scale (SIMS) was administered to assess situational motivation. A 2 (test type) \times 2 (choice) \times 2 (gender) MANOVA was used to test for significant differences on each of the four SIMS indices. Significant test type and gender and a significant test type by gender interaction were found. These findings suggest practitioners should use criterion-referenced, formative skill tests especially when teaching girls in physical education.

Keywords: curriculum, assessment, motivation, adolescent

Recent national health trends such as the “obesity epidemic” have turned the attention of lawmakers, the allied health professions, and education to solution seeking for this and other lifestyle related maladies. Concern over lifestyle related issues has even reached the Whitehouse with First Lady Obama’s recent initiative to combat childhood obesity (Executive Office of the President of the U.S., 2010). Early warnings of lifestyle diseases led to a stark list of risk factors outlined in the 1996 Surgeon General’s report (U.S. Department of Health and Human Services, 1996). In recognition of its common impact on not only obesity but many other lifestyle related diseases, topping that list was physical inactivity. It is also generally accepted that preventative measures and early intervention would be the wisest and most cost effective means to address these lifestyle related diseases.

To this end, greater scrutiny has been focused on the physical activity patterns of youth and what role physical education (PE) in the public school system can and should play in means of prevention and early intervention. Since nearly all youth

Johnson is with Boise State University—Kinesiology, Boise, ID. Prusak, Pennington, and Wilkinson are with Brigham Young University—Teacher Education, Provo, UT.

in the US and Canada are serviced by the public school system, it stands to reason that PE is considered an essential player in this national effort.

Changing the physical activity habits of an entire nation is most certainly a daunting task, one that has been likened to the battle against tobacco use (Blades, 2009). It will be a long and difficult battle that will most certainly rely heavily on the education system to instill within our youth the knowledge, skills, and dispositions toward making appropriate healthy lifestyle decisions. The disposition or *motivation* to first acquire and then act on such knowledge and to engage such skills then becomes key in the successful pursuit of healthier lifestyles.

In an effort to understand the motivational effects of PE curricula and common instructional practices researchers have used a variety of theoretical frameworks including self-determination theory (SDT; Ntoumanis, 2001; Prusak, Treasure, Darst, & Pangrazi, 2004; Standage, Duda, & Ntoumanis, 2003). For example, previous studies have found that when teachers use methodologies specifically intended to increase student perceptions of choice (Prusak et al., 2004; Ward, Wilkinson, Graser, & Prusak, 2008), perceived competence (Lee, Fredenberg, Belcher, & Cleveland, 1999), and sociality (Prusak & Darst, 2002; Reeve & Sikenius, 1994), that such practices have a positive effect on situational motivation in PE. Self-determination theory provides a particularly useful theory in that it proposes that life-long, motivated behaviors (such as physical activity patterns) are shaped and influenced by the collective and formative success in situation- and context-specific experiences, such as those in PE (Guay, Mageau, & Vallerand, 2003).

Using the tenets of SDT, Vallerand (1997, 2001) proposed a *Hierarchical Model of Intrinsic and Extrinsic Motivation* (HMIEM) that supports the notion that motivation occurs on three levels of generality, namely, situational (the activities of the present), contextual (usual motivated state toward a context, such as sport, school, or PE), and global (an omnibus personality trait toward life). Global is the most stable followed by contextual, however, situation-specific motivation can be highly changeable (i.e., situation-dependent). Vallerand's HMIEM asserts that motivational states experienced at one level can exert either a top-down or bottom-up effect on adjacent levels of generality (Guay, Mageau, & Vallerand, 2003; Prusak et al., 2004; Standage & Treasure, 2002). Thus, the cumulative effect of experiences in a given situation (such as PE activities) can affect one's physical activity contextual motivation. The idea that what we do in PE can actually have an effect on how one behaves outside of PE on up to one's life choices underscores the importance of identifying and employing practices that do so.

Self-determination theory also proposes a multidimensional perspective that distinguishes between different motives for engaging in any particular behavior including intrinsic, extrinsic, and amotivation (the absence of motives). Further, SDT posits a simplex ordering of these dimensions along a continuum with intrinsic and amotivation at distal ends and each is, in turn, multidimensional. Persons exhibiting amotivation indices are less self-determined and those who exhibit intrinsic motivation profiles are more self-determined. For the most part, however, both intrinsic and amotivation have been treated as unitary constructs because of the high degree of correlation between their respective dimensions (Green-Demers, Legault, Pelletier, & Pelletier, 2008; Standage et al., 2003). Specifically, intrinsic

motivation is characterized by (a) intrinsic motivation to know, (b) to experience stimulation, and (c) toward accomplishment. Amotivation is characterized by (a) deficient ability beliefs, (b) deficient effort beliefs, (c) insufficient academic values, and (d) unappealing characteristics of tasks. For the purposes of this paper, we will treat both as unitary constructs.

Extrinsic motivation can be assessed on four dimensions: external regulation (least autonomous form of regulation resulting from demands, rewards, or coercion); introjected regulation (engagement is externally controlled, i.e., compliance, but internally localized by a sense of guilt or pressure tension); identified regulation (a relatively autonomous acceptance of personal value placed on the activity); and integrated regulation (fully internalized motives that are decidedly self-determined). External regulation and identified regulation were measured in this study. Reasons for doing so include (a) the specificity of the measurement instrument and (b) the process of internalization, (i.e., the effectual shift from external to fully autonomous regulation) that occurs as one moves along the continuum from amotivation toward intrinsic motivation.

Lastly, a brief discussion of the antecedents of motivated behaviors identified in SDT is warranted. Environments, educational or other, that increasingly satisfy three basic needs (autonomy, competence and relatedness) will result in higher levels of self-determined motivation. Autonomy results from having a sense of having choices, or control over ones options (i.e., the locus of control lies mainly within the student). Competence results from a sense of “I can do this” or high belief for a successful outcome. Relatedness results from a sense of being able to successfully engage in the social milieu of the environment. Teachers who manipulate the learning environment to provide high levels of each of these basic needs will experience positive increases in student self-determination (Prusak et al., 2004; Prusak & Darst, 2002; Ward et al., 2008).

While more and more research on the motivational processes that exist in PE has resulted in greater understanding of the motivation effects associated with physical activity behaviors, there is still much to be done. In particular, little is known about the effects of assessment practices on the motivation of PE students. Assessments are often tied to grades and perceptions of success are heavily influenced by test outcomes. Therefore, the manner in which students are assessed in PE (specifically skill testing) should be studied to determine if manipulating testing procedures could result in increased ability to meet basic environmental needs of autonomy, competence and relatedness resulting in increased motivation.

Typically, skill testing is done at the end of a unit of instruction as a summative assessment. Standardized skills tests such as those created by the American Alliance of Health, Physical Education, Recreation and Dance (AAHPERD; Strand & Wilson, 1993) are normative-based skills tests that rank individual performance to that of a similar population. Norm referencing (also referred to as grading on the curve) results in grades being assigned according to how well someone compared with the norm. Grade breakdowns are accomplished by assigning an A-grade to those who are two standard deviations (*SD*) at or above the norm, a B-grade for one *SD* above, a C-grade for being average, and so on. These standardized skill tests are frequently not well received by teachers nor students who perceive them as overly harsh.

Prusak (2005) provides a more thorough explanation of the motivational impact of skill testing procedures and offers an alternative skill testing format—the *Programmed Practice Sheet* (PPS). The PPS is designed with the intent to increase student autonomy, competence, and relatedness. It is designed to meet the normative standards in the AAHPERD skill tests but is presented in a criterion referenced (Lund & Kirk, 2002) standards format. It presents students with choices of skill difficulty level, is formative in that it allows for repeated repetition and refinement of the skills until competence is achieved, and promotes cooperative work with fellow students.

The idea of offering choices to enhance student motivation in PE is well documented (Bryan & Solmon, 2007; Kilpatrick, Hebert, & Jacobsen, 2002; Prusak et al., 2004; Prusak & Darst, 2002; Ward et al., 2008). Commonly suggested methods to infuse choice into the PE curricula include but are not limited to (a) offering different pieces of equipment (i.e., larger or smaller balls, heavier or lighter rackets; Mosston & Ashworth, 2002), (b) providing various task/skill difficulty levels where rules for performance have been modified (i.e., complete 10 consecutive passes with a partner standing 10, 20, or 30 yards away; Mosston & Ashworth, 2002), and (c) allowing options of an activity according to a specific theme or focus (i.e., fitness, social, or competitive emphases; Prusak & Darst, 2002). This study is concerned, in part, with the impact of having or not having choices of skill difficulty level during skill testing on student situational motivation in PE.

Previous research in PE has shown gender differences across various motivation-related constructs with boys consistently demonstrating higher levels of enjoyment, expectancy-related beliefs, perceived competence, physical self-perceptions, and expectations for success than girls (Lee et al., 1999; Morgan et al., 2003; Satina, Solmon, Cothran, Loftus, & Stockin-Davidson, 1998; Xiang, McBride, Guan, & Solmon, 2003). However, data related to gender differences in situational motivation during PE are limited, especially in relation to skill testing.

The purpose of this study was to examine the effects of (a) skill test type (norm or peer-referenced, summative vs. criterion- or self-referenced, formative), (b) choice availability, and (c) gender on the situational motivation profiles of adolescent boys and girls during skill testing. It was hypothesized that the following participants would demonstrate higher levels of situational motivation: those receiving the criterion- or self-referenced, formative skill test; those receiving choices of skill difficulty level during skill testing; and boys.

Methods

Participants

Participants were male ($N = 272$) and female ($N = 235$) PE students 12–16 years of age ($M = 13.87$; $SD = 0.94$) in grades 7 ($N = 128$), 8 ($N = 175$), and 9 ($N = 204$) who attended a junior high school located in a large metropolitan area in the western U.S. Participants were predominately non-Hispanic white and came from middle-class families. Each participant was enrolled in one of 16 in-tact PE classes (14 single-gender; two coed) taught by one of three certified PE instructors (two female, one male) with one to nine years of teaching experience. A university Institutional Review Board granted permission to conduct this study, and a parent/guardian of each participant provided informed consent before data collection.

Instruments

AAHPERD Football Skill Test. The 6-item AAHPERD football skill test (Strand & Wilson, 1993), which is recommended for junior-high and high school aged students, was administered in this study. It included the following items: forward pass for distance, forward pass for accuracy, catching the forward pass, punting for distance, 50-yard dash, and the ball-changing zigzag run. Face validity has been declared for each of the six items and reliability scores of ≥ 0.80 and ≥ 0.70 were established for distance and accuracy or form events, respectively. Each of the six items was administered exactly as described by Strand and Wilson (1993). The test was administered on the last day of the flag football unit of instruction. The test was not considered complete unless students completed all six items and obtained a corresponding percentile score for each of the six items. Grades were allocated to participants based upon their percentile scores.

This test was administered in two formats: one that provided choices and one that did not. The one with choices had three levels of difficulty for participants to choose from, and the one without choices typically required participants to complete the intermediate level (Table 1). This was considered a peer- or norm-referenced, summative test because performance was compared with that of one's peers (i.e., scores determined one's age-related percentile and ultimately one's grade) and it was administered once at the end of the football unit of instruction. Earning a "good grade" on this skill test was quite difficult because (a) students received only one chance to perform (i.e., a student might have a "bad day") and (b) their performances were compared with their age-related peers (i.e., not everyone could get a high grade).

Programmed Practice Sheet Skill Test. A PPS was designed and created for use in this study as the self- or criterion-referenced, formative skill test. The PPS is a rubric or task sheet that lists the desired learning outcomes and prescribes a systematic program-of-practice toward their achievement. The PPS created for this study (see Appendix) included derived items from the six football skills part of the AAHPERD Football Skill Test. As a result of patterning the PPS items after the AAHPERD Football Skill Test, face validity and reliability of the PPS were assumed.

Table 1 AAHPERD Skill Test Items for the Choice and No-Choice Groups

Skill Test Item	Choices
Forward Pass for Distance	Small, <i>intermediate</i> -, or official-sized football
Forward Pass for Accuracy	Small, <i>intermediate</i> -, or official-sized football
Catching the Forward Pass	Distances of 20, 30, or 40 feet
Punting for Distance	Soccer ball, <i>intermediate</i> -, or official-sized football
50-Yard Dash	Distances of 30, 40, or 50 yards
Ball Changing Zig-Zag Run	Distances of 8, 10, or 12 feet between cones

Note. Italicized choices were provided to the no-choice group

Two PPSs were created and used during this study, one offering choices and the other without choices. This was considered a self- or criterion-referenced, formative test because participants were given preset skill criterion standards and multiple testing opportunities throughout the football unit of instruction to achieve those standards. Students who met the standards were given full credit or points. In other words, students could earn a “good grade” with persistence and effort.

Situational Intrinsic Motivation Scale (SIMS). The 16-item SIMS was administered to students immediately following the skill testing protocol. Specifically, the SIMS measures the following motivational states of SDT: amotivation (AM), external regulation (ER), identified regulation (IR), and intrinsic motivation (IM). The SIMS stem question for this study stated, “Why are you currently engaged in this skill test?” Students responded using a 7-point Likert scale to items like the following: (a) “because I think this activity is pleasant,” (b) “because I think this activity is good for myself,” (c) “because it is something I have to do,” (d) “I do this activity but I am not sure it is worth it.” The SIMS has been shown to be a valid and reliable measure of the above mentioned motivational states (Guay, Vallerand, & Blanchard, 2000; Standage & Treasure, 2002). In addition, research utilizing the SIMS in field settings has consistently supported the simplex-ordered pattern of the self-determination continuum with intrinsic motivation and amotivation on distal ends and the corresponding forms of extrinsic motivation in the middle. In other words, correlations between SIMS subscales have been in the anticipated directions and of expected magnitudes. For further information about the use of the SIMS instrument in field settings see Prusak et al. (2004) and Ward et al. (2008).

A mean score (derived from four of the 16 items on the SIMS) can be calculated for each of the four motivational states (AM, ER, IR, & IM). Higher mean scores denote higher levels of that corresponding motivational state. A Self-Determination Index (SDI) score, or level of self-determination, can be computed using data from the SIMS instrument. To do so, the mean of each motivational state is needed in the following formula: $[SDI = +2(IM) + 1(IR) - 1(ER) - 2(AM)]$. The SDI is easy to interpret—higher scores equate to higher levels of situational motivation. The SDI score has been described and used in field settings previously by Pelletier et al. (1995) and Prusak et al. (2004), respectively.

Procedures

One week before the beginning of the flag football unit of instruction, a 90-min orientation meeting was conducted by the research team to teach the PE instructors about the skill testing protocol. Instructors were given all needed skill testing materials such as necessary instructions, equipment, and scoring sheets, and were taught to incorporate the assigned skill testing protocol into the unit where applicable. During the orientation meeting, all three of the instructors personally completed portions of the skill tests so as to develop adequate familiarity. Instructors implemented the skill testing protocol into their curriculums during or at the end of a three-week flag football unit that included 8 lessons. Classes met every other day for 80 min. Flag football was one of many team and individual sports/activities taught as part of the school PE curriculum.

During lesson 1 of the flag football unit, students in each class were introduced, by their instructors, to the skill test assigned to their class and also how their performance on the skill test would be used to allocate skill grades for the flag football unit. During lessons 1–7, all groups received direct, formal instruction (i.e., explanation, demonstrations, cues, etc.) from their instructors regarding how to perform flag football skills (throwing for accuracy, throwing for distance, catching, punting for distance, etc.). All groups were given time in class to practice the skills. During practice time, instructors of students receiving the AAHPERD skill test provided skill-related feedback and reminded participants about the skill test that would be administered on the last day of the unit. Instructors of students assigned to receive the PPS skill test also provided skill-related feedback and encouraged them to exert their best efforts in completing the items listed thereon. On the last day of the unit, students received and completed the AAHPERD skill test or finished and turned in the PPS. All students knew their corresponding skill test grade before turning in their scoring sheets. Immediately following the completion of their associated skill test, a research team member explained, distributed, and collected a SIMS instrument to and from each student. If a student was absent the day a skill test was administered, he or she completed it the next available class period. It is important to note that participants had never experienced any of the skill tests administered in this study previously in their physical education classes.

Design and Data Analyses

Each of the 16 classes was randomly assigned to one of four groups (4 classes in each group): the AAHPERD test (with or without choices) or the PPS (with or without choices). Although the treatment was administered at the class level, analyzing data at the student-level was deemed appropriate because skill testing was conducted individually or in pairs, and it was unclear if mere class membership would cause class-level effects.

A manipulation check, or method of identifying test type, was administered to a panel of three experts from a large university and the three physical education instructors before the intervention began. It consisted of panel members and instructors identifying the correct test type by matching the name of the skill test (i.e., peer- or norm-referenced, summative with and without choices; self- or criterion-referenced, formative with and without choices) with the corresponding instruction and scoring sheets. In addition, a manipulation check was administered to students. On a sheet of paper, students were asked “what type of test do you think you had” and were asked to circle one of the following choices: (a) “I had choices during this skill test, and I was given credit for the work I did *during the unit*,” (b) “I did not have choices during this skill test, and I was given credit for the work I did *during the unit*,” (c) “I had choices during this skill test, and I was tested on how well I did at the *end of the unit*,” and (d) “I did not have choices during this skill test, and I was tested on how well I did at the *end of the unit*.” Student responses were scored from zero to three and were entered into the data set; a discriminant analysis was then calculated to determine the effectiveness of the manipulation check.

Cronbach’s alpha scores and pearson correlation coefficients were computed to determine the internal consistency/reliability and the simplex-ordered pattern of the SIMS instrument. Descriptive statistics (means, standard deviations) were calculated

for each of the SIMS indices (AM, ER, IR, IM) and the SDI score. A 2 (test type) \times 2 (choice) \times 2 (gender) MANOVA was used to examine between group differences and interactions. Significance was established a priori at $p \leq .05$. Effect sizes were computed when significant differences existed between groups (Cohen, 1990). All statistical analyses were calculated using SPSS version 17.0 (SPSS, Inc., Chicago, IL.).

Results

Manipulation Check

Both the panel of experts and the instructors correctly identified (100%) each skill test type. Seventy-eight percent of the students accurately selected the test type they received. Results of the discriminant analysis showed that student responses were significantly different when asked what test type they had ($\chi^2(1, N = 482) = 325.48; p < .001$) demonstrating they could discern their corresponding test type supporting the success of the intended manipulation.

Reliability and Internal Consistency

Cronbach's alpha for the four SIMS indices revealed adequate internal consistency values ranging from 0.81 to 0.85. Reliability scores are considered acceptable when values are ≥ 0.7 (see Table 2, α values on diagonal).

Table 2 Correlations and Internal Consistency

SIMS Subscales	1	2	3	4
1. Amotivation	(.83)	.58*	-.48*	-.56*
2. External Regulation		(.82)	-.34*	-.39*
3. Identified Regulation			(.81)	.81*
4. Intrinsic Motivation				(.85)

Note. The pattern of strongest positive correlations between AM/ER and IR/IM to strongest negative correlations between IM, ER, and AM supports the simplex pattern of the SIMS instrument. Values in parentheses are Cronbach's alpha coefficients.

* $p < .01$.

Simplex Pattern of SIMS

Pearson correlation coefficients for all pairs of the four SIMS indices were calculated (see Table 2, above diagonal). All correlations were found to be significant at the $p < .01$ level and in the anticipated direction. This supports the simplex pattern of the SIMS instrument (see note in Table 2) because these correlations maintain the idea that AM, ER, IR and IM lie along a motivation continuum with AM and IM on distal ends.

Situational Motivation

Descriptive statistics (means, standard deviations, and effect sizes) are presented in Table 3. Results from the MANOVA indicated a main effect for test type (AM; $F(1, 506) = 4.558; p < .05$; ER; $F(1, 506) = 4.151; p < .05$; IR; $F(1, 506) = 6.119; p < .05$;

IM; $F(1, 506) = 4.451$; $p < .05$; SDI; $F(1, 506) = 7.256$; $p < .01$) and gender (AM; $F(1, 506) = 13.762$; $p < .001$; ER; $F(1, 506) = 43.251$; $p < .001$; IR; $F(1, 506) = 29.022$; $p < .001$; IM; $F(1, 506) = 35.779$; $p < .001$; SDI; $F(1, 506) = 42.074$; $p < .001$) across each of the SIMS indices and the SDI score. Significant test type by gender interactions were also found (AM; $F(1, 506) = 4.363$; $p < .05$; ER; $F(1, 506) = 9.388$; $p < .01$; IR; $F(1, 506) = 21.325$; $p < .001$; IM; $F(1, 506) = 11.246$; $p < .01$; SDI; $F(1, 506) = 14.583$; $p < .001$). These analyses indicated that (a) boys reported higher levels of self-determination during skill testing than girls and (b) girls were more affected by test type than boys. Boys reported very similar SIMS scores on both test types whereas girls receiving the AAHPERD skill test reported lower self-determination scores than those receiving the PPS skill test across each of the SIMS indices and the SDI score (Table 4). For example, girls receiving the AAHPERD skill test ($M = 3.15$; $SD = 1.48$) reported higher levels of amotivation than girls receiving the PPS ($M = 2.55$; $SD = 1.53$; Table 4 contains descriptive statistics by test type and gender).

Discussion

This study examined the effects of skill test type, choice availability, and gender on the situational motivation of adolescent youth in PE class. The general findings of this study revealed that (a) skill test type (receiving the PPS) and gender (being

Table 4 Means and Standard Deviations (AM, ER, IR, IM, SDI) for Test Type by Gender

AAHPERD	PPS		M	SD
	M	SD		
AM				
Girls	3.15*	1.48	2.55	1.53
Boys	2.32	1.63	2.27	1.57
ER				
Girls	4.45**	1.45	3.75	1.60
Boys	3.10	1.50	3.22	1.56
IR				
Girls	2.28***	1.44	3.13	1.45
Boys	3.56	1.32	3.32	1.45
IM				
Girls	2.18**	1.56	2.86	1.55
Boys	3.45	1.49	3.29	1.40
SDI				
Girls	-4.09***	7.42	.002	7.67
Boys	2.72	7.28	2.13	6.86

Note. Asterisks indicate a significant overall test type by gender interaction

* $p < .05$. ** $p < .01$. *** $p < .001$.

male) led to greater levels of situational motivation compared with receiving the AAHPERD skill test and being female (see Table 3) and (b) girls were more affected by test type than boys (see Table 4).

Skill Test Type

Deci and Ryan (1987) argue when the “locus of control” or potential for success in an activity is outside one’s control or when one perceives a limited possibility of success, motivation decreases. This likely happened to students who received the AAHPERD skill test. Students receiving this skill test likely perceived their chances of success (i.e., demonstrating competence, getting a good grade), even if they exerted lots of effort, to be outside their control and as a result reported relatively high levels of AM and ER and low levels of IR and IM. In contrast, participants who received the PPS skill test likely perceived the “locus of control” to be inside of themselves or somewhat under their control because the skills could be practiced sufficiently until criterion standards were achieved. As a result, their reported levels of AM and ER were lower and levels of IR and IM higher than those receiving the AAHPERD skill test. From a motivational perspective, practitioners who administer skill tests should employ self-referenced, formative tests, like the PPS, rather than norm-referenced, summative tests like the AAHPERD skill test.

Choices and Skill Testing

A tenet of SDT (Deci & Ryan, 1985) is that students will experience higher levels of motivation when they feel autonomous (i.e., “I have choices”) or are given some decision-making opportunities. The decision to offer choices of skill difficulty levels during skill testing in this study was made because previous studies have shown higher levels of self-determined participation when choices of activities were offered (Prusak et al., 2004; Ward et al., 2008). In addition, Treasure and Roberts (1995) have argued that practitioners who share decision-making with their students, such as providing various skill difficulty levels, are more likely to foster adaptive motivational behaviors in their students. Interestingly, this study did not show significant differences in situational motivation between the choice and no choice skill testing groups. This suggests that offering choices during skill testing may not override the motivational impact of the type of skill test administered or perhaps the choices offered were not clearly distinguishable from each other to elicit changes in situational motivation. No doubt further examination is needed on this subject.

Gender and Skill Testing

In this study, boys exhibited higher levels of IM and IR and lower levels of ER and AM and an overall higher SDI score than girls during skill testing (see Table 3). This finding is not surprising since boys have also shown higher levels of perceived competence, expectancy-related beliefs, and expectations for success in PE (Morgan et al., 2003; Satina, Solmon, Cothran, Loftus, & Stockin-Davidson, 1998; Xiang, McBride, Guan, & Solmon, 2003). However, the type of activity selected (i.e., football) for this study likely favored boys’ interests compared with girls and probably contributed to the gender difference. Research has shown that adolescents

perceive team sports, like flag football, to be for boys and creative activities such as dance and gymnastics to be for girls (Clifton & Gill, 1994). Future studies should explore the effects of different types of skill testing within other activities (i.e., creative dance, gymnastics, tennis, etc) to further understand gender differences in situational motivation during skill testing.

This study showed a significant test type by gender interaction across each of the SIMS indices and the SDI score highlighting that girls and boys were affected differently by the type of skill test administered. Specifically, girls reported higher levels of IM and IR, lower levels of AM and ER, and an overall higher SDI score for the PPS skill test than the AAHPERD skill test. Boys, on the other hand, reported very similar situational motivation scores on both skill tests.

Possible reasons why girls were more affected by skill test type might include: first, girls likely perceived the “locus of control” to be outside of themselves. Perhaps because football was the selected sport, girls receiving the AAHPERD skill test possibly perceived no matter how much effort they exerted, their chances for success (i.e., get a good grade) were outside of their control. In contrast, the “locus of control” for girls receiving the PPS likely became more internalized leading to feelings of “I can do this” (confidence) because they could take the skill test multiple times until they met the specified standards of performance. In other words, girls receiving the PPS probably felt some control in dictating the skill grade they could achieve. Second, in physical education, girls have shown a stronger preference for cooperative rather than competitive activities (Prusak & Darst, 2002). The AAHPERD skill test, primarily due to its norm- or peer-referenced criteria for success, likely created a more competitive than cooperative environment thus reducing girls’ situational motivation.

Research has shown that adolescent girls are not as physically active as adolescent boys and children (Le Masurier et al., 2005; Trost, Pate, & Sallis, 2002). However, frequent and repeatedly positive situational experiences (of which skill testing in several sports/activities over the course of a school year or years is part) may exert bottom-up effects and eventually lead to changes in one’s lifestyle choices (Deci & Ryan, 1985; Vallerand, 2001). The findings of this study suggest practitioners should use self-referenced, formative skill tests, like the PPS, when teaching adolescent girls to promote higher levels of situational motivation. Less clear are the effects of test type on boys’ situational motivation warranting further investigation. It is interesting to note that boys receiving the AAHPERD skill test reported slightly higher IM and IR scores and similar AM and ER scores than boys receiving the PPS.

Limitations & Conclusions

This study was not without limitations. First, one of the PE instructors in this study was the primary researcher who knew the objectives of the study. This instructor taught all of the single-gender boys PE classes (seven). Second, the unit of analysis was a notable limitation that must be remembered when interpreting the results of this study. Since the treatment was administered at the class level, statistical analyses should also have been conducted at the class level (Silverman & Solmon, 1998). The decision to analyze data at the individual level was the only plausible option since including enough clusters (classes) was overly daunting given the available resources to conduct this study. However, the results of this study provide preliminary evidence regarding the impact of

skill test type and gender on student situational motivation levels. Future studies should address the unit of analysis issue by planning for additional clusters and thus enabling other statistical analyses (i.e., hierarchical linear modeling). Third, the intervention was relatively short and included skill testing for only one sport/activity—football—which may have contributed to how boys and girls experienced the skill testing. Fourth, both test types had two subfactors—being norm- or criterion-referenced and being summative or formative. As a result, it is difficult to say which subfactor, or whether a combination of both, led to the differences in situational motivation.

Despite these limitations, this study provides some support for using criterion- or self-referenced, formative skill tests in PE. Generally speaking, girls were more likely to demonstrate higher levels of situational motivation when administered such a test compared with a norm- or peer-referenced, summative test. It has been observed that adolescent girls prefer noncompetitive activities which could have been a reason why the norm- or peer-referenced, summative skill test did not appeal to them. Future research studies should (a) evaluate the effects of the types of skill testing over an entire semester or year of PE and across a variety of activities and/or instructional units; (b) determine if and how these types of skill tests affect student perceptions of the motivational climate; and (c) include more physical education classes to enable analysis at the class level.

References

- Blades, L. (2009, June). *Physical activity, nutrition, and obesity program*. Paper presented at the meeting of Utah Department of Health Conference, Draper, UT.
- Bryan, C.L., & Solmon, M.A. (2007). Self-determination in physical education: Designing class environments to promote active lifestyles. *Journal of Teaching in Physical Education, 26*, 260–278.
- Clifton, R.T., & Gill, D.L. (1994). Gender differences in self-confidence on a feminine-typed task. *Journal of Sport & Exercise Psychology, 16*, 150–162.
- Cohen, J. (1990). Things I have learned (so far). *The American Psychologist, 45*, 1304–1313.
- Deci, E.L., & Ryan, R.M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum Press.
- Deci, E.L., & Ryan, R.M. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology, 53*, 1024–1037.
- Executive Office of the President of the U.S. White House Task Force on Childhood Obesity: Report to the President. (2010). Solving the problem of childhood obesity within a generation. Retrieved from http://www.letsmove.gov/pdf/TaskForce_on_Childhood_Obesity_May2010_FullReport.pdf
- Green-Demers, I., Legault, L., Pelletier, D., & Pelletier, L.G. (2008). Factorial invariance of the academic amotivation inventory across gender and grade in a sample of Canadian high school students. *Educational and Psychological Measurement, 68*, 862–880.
- Guay, F., Mageau, G.A., & Vallerand, R.J. (2003). On the hierarchical structure of self-determined motivation: A test of the top-down, bottom-up, reciprocal, and horizontal effects. *Personality and Social Psychology Bulletin, 29*, 992–1004.
- Guay, F., Vallerand, R.J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The situational motivation scale (SIMS). *Motivation and Emotion, 24*, 175–213.
- Kilpatrick, M., Hebert, E., & Jacobsen, D. (2002). Physical activity motivation: A practitioner's guide to self-determination theory. *Journal of Physical Education, Recreation & Dance, 73*(4), 36–41.

- Lee, A.M., Fredenberg, K., Belcher, D., & Cleveland, N. (1999). Gender differences in children's conceptions of competence and motivation in physical education. *Sport Education and Society*, 4, 161–174.
- Le Masurier, G.C., Beighle, A., Corbin, C.B., Darst, P.W., Morgan, C., Pangrazi, R.P., et al. (2005). Pedometer-determined physical activity levels of youth. *Journal of Physical Activity and Health*, 2, 159–168.
- Lund, J.L., & Kirk, M.F. (2002). *Performance-based assessment for middle and high school physical education*. Champaign, IL: Human Kinetics.
- Morgan, C.F., McKenzie, T.L., Sallis, J.F., Broyles, S.L., Zive, M.M., & Nader, P.R. (2003). Personal, social, and environmental correlates of physical activity in a bi-ethnic sample of adolescents. *Pediatric Exercise Science*, 15, 288–301.
- Mosston, M., & Ashworth, S. (2002). *Teaching physical education* (5th ed.). San Francisco, CA: Benjamin Cummings.
- Ntoumanis, N. (2001). A self-determination approach to the understanding of motivation in physical education. *The British Journal of Educational Psychology*, 71, 225–242.
- Pelletier, L.G., Fortier, M.S., Vallerand, R.J., Tuson, K.M., Briere, N.M., & Blais, M.R. (1995). Towards a new measure of intrinsic motivation, extrinsic motivation, and amotivation in sports: The Sport Motivation Scale. *Journal of Sport and Exercise Psychology*, 17, 35–53.
- Prusak, K.A. (2005). Assessing students in the task involved motivational climate. *Teaching Elementary Physical Education*, 16, 11–17.
- Prusak, K.A., & Darst, P.W. (2002). Effects of types of walking activities on actual choices by adolescent female physical education students. *Journal of Teaching in Physical Education*, 21, 230–241.
- Prusak, K.A., Treasure, D.C., Darst, P.W., & Pangrazi, R.P. (2004). The effects of choice on the motivation of adolescent girls in physical education. *Journal of Teaching in Physical Education*, 23, 19–29.
- Reeve, J., & Sikenius, B. (1994). Development and validation of a brief measure of the three psychological needs underlying intrinsic motivation. *Educational and Psychological Measurement*, 54(2), 506–515.
- Satina, B., Solmon, M.A., Cothran, D.J., Loftus, S.J., & Stockin-Davidson, K. (1998). Patriarchal consciousness: Middle school students' and teachers' perspectives of motivational practices. *Sport Education and Society*, 3, 181–200.
- Silverman, S., & Solmon, M. (1998). The unit of analysis in field research: Issues and approaches to design and data analysis. *Journal of Teaching in Physical Education*, 17, 270–284.
- Standage, M., Duda, J.L., & Ntoumanis, N. (2003). Predicting motivational regulations in physical education: the interplay between dispositional goal orientations, motivational climate and perceived competence. *Journal of Sports Sciences*, 21, 631–647.
- Standage, M., & Treasure, D.C. (2002). Relationship among achievement goal orientations and multidimensional situational motivation in physical education. *The British Journal of Educational Psychology*, 72, 87–103.
- Strand, B.N., & Wilson, R. (1993). *Assessing sport skills*. Champaign, IL: Human Kinetics.
- Treasure, D.C., & Roberts, G.C. (1995). Applications of achievement goal theory to physical education: Implications for enhancing motivation. *Quest*, 47, 475–489.
- Trost, S.G., Pate, R.R., & Sallis, J.F. (2002). Age and gender differences in objectively measured physical activity in youth. *Medicine and Science in Sports and Exercise*, 34, 350–355.
- U.S. Department of Health and Human Services. (1996). *Physical Activity and Health: A Report of the Surgeon General*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion.
- Vallerand, R.J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. In M.P. Zanna (Ed.), *Experimental Social Psychology* (Vol. 29, pp. 271–361). Toronto: Academic Press.

- Vallerand, R.J. (2001). A hierarchical model of intrinsic and extrinsic motivation in sport and exercise. In G. Roberts (Ed.), *Advances in Motivation in Sport and Exercise* (pp. 263–319). Champaign, IL: Human Kinetics.
- Ward, J.W., Wilkinson, C., Vincent Graser, S., & Prusak, K.A. (2008). Effects of choice on student motivation and physical activity behavior in physical education. *Journal of Teaching in Physical Education*, 27, 385–398.
- Xiang, P., McBride, R., Guan, J., & Solmon, M. (2003). Children's motivation in elementary physical education: An expectancy-value model of achievement choice. *Research Quarterly for Exercise and Sport*, 74, 25–35.

Appendix

Flag Football Programmed Practice Sheet (PPS)

Level A Easiest	Level B Medium	Level C Advanced
Throw and catch 10 passes with a partner 15 yards apart. The receiver can take no more than one step to catch the ball. Repeat three times and initial in space each time 1. ____ 2. ____ 3. ____	Throw and catch 10 passes with a partner 15 yards apart. The receiver must keep one foot planted. Repeat three times and initial in space each time 1. ____ 2. ____ 3. ____	Throw and catch 10 passes with a partner running across 15 yards in front of the passer. Repeat twice and initial in space each time 1. ____ 2. ____
Throw and catch 25 consecutive passes with a partner 15 yards apart. Do this three times and initial each time completed. 1. ____ 2. ____ 3. ____	Throw and catch 25 consecutive passes with a partner 20 yards apart. Do this three times and initial each time completed 1. ____ 2. ____ 3. ____	Throw and catch 25 consecutive passes with a partner 25 yards apart. Do this twice and initial each time it is completed 1. ____ 2. ____
Catch 10 out-routes run at 10 yards. Dropped balls or bad passes can be repeated as often as necessary. Repeat three times and initial. 1. ____ 2. ____ 3. ____	Catch 10 out-routes run at 15 yards. Dropped balls or bad passes can be repeated as often as necessary. Repeat three times and initial. 1. ____ 2. ____ 3. ____	Catch 10 out-routes run at 15 yards. Must use thumbs-in technique. Dropped balls or bad passes can be repeated as often as necessary. Repeat three times and initial. 1. ____ 2. ____ 3. ____
Run, carrying the ball, the 25 yard dash three times and record your fastest of three times in the space. Repeat three times. 1. ____ 2. ____ 3. ____	Run, carrying the ball, the 50 yard dash two times and record your fastest of two times in the space. Repeat three times. 1. ____ 2. ____ 3. ____	Run, carrying the ball, the 100 yard dash one time and record your time in the space. Repeat three times. 1. ____ 2. ____ 3. ____
Run the ball changing Zig-Zag with single-file cones, three times and record the fastest of your three times in the space. Repeat three times. 1. ____ 2. ____ 3. ____	Run the ball changing Zig-Zag with staggered cones, three times and record the fastest of your three times in the space. Repeat three times. 1. ____ 2. ____ 3. ____	Run the ball changing Zig-Zag with double-file cones, three times and record the fastest of your three times in the space. Repeat three times. 1. ____ 2. ____ 3. ____
Punt the ball for distance three times. Record the best of three kicks in the space. Repeat three times. 1. ____ 2. ____ 3. ____	Punt the ball for accuracy three times through the uprights from the ten-yard line. Repeat three times. 1. ____ 2. ____ 3. ____	Punt the ball for accuracy three times through the uprights from the 15-yard line. Repeat twice 1. ____ 2. ____

Note. Level B was administered to the no-choice PPS group.