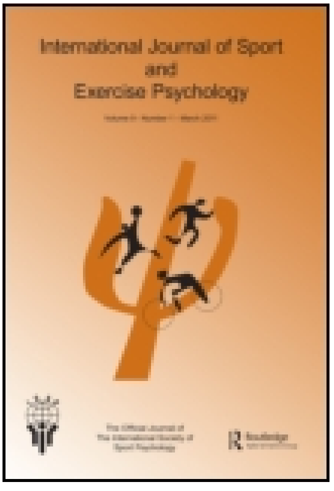

International Journal of Sport and Exercise Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rijs20

Motivational constructs in Greek physical education classes: Factor structure, gender and age effects in a nationally representative longitudinal sample

Herbert W. Marsh [a e] , Athanasios Papaioannou [b] , Andrew J. Martin [c] & Yannis Theodorakis [d]

[a] University of Western Sydney

[b] Democritus University of Thrace

[c] University of Western Sydney

[d] University of Thessaly

[e] Department of Educational Studies , University of Oxford , 15 Norhams Gardens, Oxford, OX2 6PY E-mail:
Published online: 28 Feb 2011.

IJSEP, 2006, 4, 121-148

Motivational Constructs in Greek Physical Education Classes: Factor Structure, Gender and Age Effects in a Nationally Representative Longitudinal Sample

[1,5]Herbert W. Marsh, [2]Athanasios Papaioannou, [3]Andrew J. Martin, and [4]Yannis Theodorakis

[1]University of Western Sydney
[2]Democritus University of Thrace
[3]University of Western Sydney
[4]University of Thessaly
[5]University of Oxford

Abstract

This study evaluates the construct validity of a profile of 11 motivational constructs (Task and Ego Orientation, Task- and Ego-involving Climates; Intrinsic Motivation Enjoyment; Intrinsic Motivation Effort; Exercise Attitudes; Exercise Intentions; Perceived Behavioral Control; Actual Exercise Behavior; and Physical Self-concept) appropriate for applied sport/exercise settings. A nationally representative sample of Greek students (2786, 50% males) from 200 physical education classes at different levels of schooling (29% upper primary, 36% middle, and 35% high school) completed the same battery of tests at the start and end of the school year. Despite the brevity of some of the measures (11 scales, 46 items) dictated by the large scale of the study, reliability estimates were mostly adequate (Md α = .82 at time 1, .86 at time 2). Confirmatory factor analyses of the combined set of time 1 and time 2 responses provided an excellent fit to the data (RMSEA = .034) and moderate test-retest correlations (.37 to .64; Md = .55). Most outcomes decreased with age. Whereas boys had higher scores on most constructs, there were no gender differences for Task Orientation and Task-involving Climate, and girls had slightly more positive attitudes toward exercise. The psychometric results and patterns of relations among constructs provide good support for the construct validity of the measures.

Key words: sport motivation, classroom climate, confirmatory factor analysis, construct validity, gender differences, test-retest correlation

Corresponding author: Herbert W. Marsh, Department of Educational Studies, University of Oxford, 15 Norhams Gardens, Oxford OX2 6PY; Email: herb.marsh@edstud.ox.ac.uk

The overarching intent of the present investigation was to evaluate evidence in support of the construct validity of a set of motivational constructs in sport and exercise settings. Increasingly, there is general agreement among sport/exercise psychology researchers and practitioners for the need to develop sport-specific instruments and to evaluate them within a construct validity framework (Marsh, 2002). In his review of sport and exercise tests, Ostrow (1990) reported substantial gains over the last 25 years, but emphasized that many tests are still "one shot assessments," lacking further development and refinement. Historically, despite recognition of the importance of developing reliable and valid measures, it has been evident that the quality of measures in sport/exercise research has been weak. Of the 175 instruments summarized in the Directory of Psychological Tests in Sport and Exercise Sciences (Ostrow, 1990), only one third had items based on a conceptual or theoretical framework, less than one in four reported factor analyses, and less than ten percent showed evidence of extensive reference support. However, in the last decade there has been substantial progress in the application of construct validation in sport/exercise psychology (e.g., see Duda, 1998; Marsh, 2002).

Consistent with this progress, the present study pursues a rigorous construct validation approach to the issue of psychological assessment and motivation in physical education settings with a large, nationally representative longitudinal sample of Greek school students (Papaioannou, Marsh, & Theodorakis, 2004). The sample composition and size make it the most substantial and high quality pool of Greek students ever examined in the context of motivation. Heightening students' motivation in school-based physical education is expected to increase youngsters' physical activity and to enhance subsequent activity later in life (Corbin, 2002; Duda, 1996). The sharp decline of physical activity in adolescence (Rowland, 1999) and the epidemic of inactivity that may ultimately result in hundreds of thousands of deaths each year in modern societies (McGinnis, 1992) suggests there is a need to elevate physical activity promotion as the main goal of contemporary physical education (Corbin, 2002; McKenzie, 2001). Taken together, the high quality and substantial sample combined with the rigorous analytical approach adopted in the study, underscore the importance of the study in contributing to current understanding of motivation in the physical education context.

The study also brings together a coherent profile of motivation-related constructs in the same study. The constructs selected for this model were based on our theorizing that there exist individual- and climate-related motivation constructs (comprising ego-involving climate, task-involving climate, ego goal orientation, task goal orientation) that are logically connected to process-related factors (encompassing enjoyment and effort), one's self-conceptions of physical capacity (physical self-concept), and a variety of planned and actual behavior factors (encompassing exercise attitudes, exercise intent, perceived behavioral control, and behavior). Hence, we approach our analysis through an integrative framework on the basis of four conceptual categories comprising (a) individual- and climate-related motivation, (b) process factors, (c) self-conceptions, and (d) behavior. The selected constructs are considered the best correlates of physical activity of children and adolescents that can be shaped by the physical education environment (Sallis, Prochaska, & Taylor, 2000). The selection and integration of constructs is also important from an applied perspective because, we argue, they reflect the broader process relevant to physical activity and enable practitioners to conduct a more focused analysis of the psychological constructs underpinning this process. Scales considered for this nation-wide study were selected on the basis of their theoretical importance, practical relevance, and broad generality over a wide range of ages of children and adolescents. Hence, an important criterion to be evaluated in the present investigation is the invariance of the factor structure over different age groups. However, because of the large number of scales to be administered on multiple occasions, the design of the instrument sought to achieve a strategic balance between psychometric rigor typically associated with longer scales (i.e., more items per scale) and brevity associated with short scales that would enhance its usefulness. Hence, we evaluate the success of relatively brief scales considered here in achieving a satisfactory level of psychometric rigor.

Although there has been considerable conceptual convergence on sport psychological constructs such as those considered here, there has also been ongoing debate about the degree of overlap between apparently distinct constructs—particularly those coming from different theoretical frameworks and primarily used by different "camps" of researchers who typically do not systematically evaluate how their measures of their constructs relate to those used by other researchers. Illustrating this concern, Marsh (1994b) factor-analyzed responses to two different sport motivation instruments and found that whereas mastery and goal scales from the two instruments were highly related and reflected a common underlying factor, the competition scale from one instrument reflected primarily a performance orientation but the competition scale from the other instrument reflected more of a task orientation than a performance or ego orientation. Based on these results he warned researchers to beware of jingle (assuming that scales with the same name reflect the same construct) and jangle (assuming that scales with different names reflect different constructs) fallacies and to pursue construct validity studies more vigorously to test interpretations of the measures. Heyman and Dweck (1992) similarly warned motivation researchers that they "need to take care that they are not measuring the same construct disguised in different scale names" (p. 243). Hence, a stronger focus on construct validation in relation to other constructs should force sport psychology researchers to evaluate more carefully the convergent and discriminant validity of their constructs. At the level of individual items, the observation that items on a given scale all load on a single factor when only one scale is considered provides no evidence that the items will not be more strongly related to different factors when items from a range of different constructs are considered in the same factor analysis. At the level of scales, the heuristic label assigned a particular collection of items posited to represent a scale is not a sufficient basis for establishing that the scale is highly related to other constructs with seemingly similar labels or that it is distinct from other constructs that have seemingly different labels.

A Construct Validation Approach

All constructs in sport/exercise psychology are hypothetical constructs and so must be validated using a construct validity approach so that their usefulness must be established by investigations of their construct validity (Marsh, 1997, 2002). From a construct validation perspective, theory, measurement, empirical research, and practice are inexorably intertwined so that the neglect of one will undermine the others. Ideally, validation is an ongoing process in which theory and practice are used to develop a measure, empirical research is used to test the theory and the measure, both the theory and the measure are revised in relation to research, new research is conducted to test these refinements, and theory and research are used to inform practice. Construct validity investigations can be classified as within-network or between-network studies (Marsh, 1997, 2002). Within-network studies explore the internal structure of a construct. They test, for example, the dimensionality of the construct and may seek to show that the construct has consistent, distinct multidimensional components. Factor analysis and related statistical procedures have been the primary tool for within-network studies. Between-network studies attempt to establish a logical, theoretically consistent pattern of relations between measures of a construct and other constructs. The resolution of at least some within-construct issues should be a logical prerequisite to conducting between-construct research, but researchers are often seduced into pursuing between-network research before they have done the hard work of developing an appropriate measure, evaluating the psychometric properties of responses, validating the structure (using item analysis, reliability, test-retest correlation, and, particularly, factor analysis), and revising their measures appropriately. Although purely exploratory factor analyses do not always result in theoretically relevant, replicable constructs, confirmatory factor analysis (CFA) encourages the development of instruments to measure specific, a priori factors—often derived from an explicit theoretical basis. Whereas Schutz and Gessaroli (1993) lamented the under-utilization of CFA in the sport sciences, its use is becoming more widespread in this field of research (e.g., Marsh, 1997, in press; Duda, 1998).

Part of the role of construct validation is to establish the relations between a particular construct and other constructs to which it was logically related. A problem facing consumers of sport and exercise measures is that research based on a particular measure is typically pursued by a particular group of researchers with a specific research agenda, whereas applied researchers are often interested in using a particular measure as one component of a more extensive inventory. Whereas a measure of a particular construct—or logically related set of constructs—may appear to have a coherent factor structure when measured in isolation, the apparently clean factor structure may not be as clean when the measures are factor analyzed as part of a more extensive battery of related constructs. The present investigation adopts a construct validation approach to the study of an a priori profile of motivation-related constructs selected specifically for their relevance in a physical activity setting. In particular, it does so through an integrative framework comprising a coherent profile of motivation measures that, we propose, provides a powerful test of the constructs under study. As described above, a rigorous analysis of constructs is one that tests their robustness in the context of other conceptually related factors.

Overview of the Constructs to Be Evaluated in the Present Investigation

The present investigation draws together a profile of motivation-related constructs in one study. We proposed a conceptually coherent profile of constructs revolving around the notion that there exist individual- and climate-related motivation constructs that are logically connected to process-related factors, self-conceptions of physical capacity, and a variety of planned and actual behavior factors. This integrative approach is an important strength of the study, and one that provides a stringent test of the constructs under examination, given that analysis of them is carried out in the context of an extensive battery and not individually. Moreover, the integration of constructs is important from an applied perspective as they reflect the broader process relevant to physical activity and enable practitioners to conduct a more focused analysis of the psychological constructs underpinning this process.

Individual- and climate-related motivation

Motivation goal theory, stemming largely from the research by Nicholls and colleagues (Duda, 2001; Duda & Nicholls, 1992; Nicholls, 1989) and others (Roberts, 1993; Marsh, Craven, Hinkley & Debus, 2003; Papaioannou & Theodorakis, 1996), has focused on two contrasting dispositional goal orientations. Central to a task orientation is attention to the processes of successfully completing or mastering tasks; development of increased competency and knowledge, the endorsement of the intrinsic value of learning as an end in itself and the belief that appropriate effort will result in better performance. Central to an ego orientation is a focus on social comparison processes in which the individual "beats" other students or attains success based on little effort, a desire to gain positive judgments and avoid negative judgments of one's competence, external evaluations of self, endorsement of the extrinsic value of performance as a means to a desired goal, and beliefs that ability is a relatively fixed attribute that cannot be altered by effort. In her review of goal orientation research in physical education and sport settings, Duda (2001) emphasized that task and ego goal orientations were dispositional (individual difference) variables that were reasonably orthogonal—not bipolar—when measured with the most widely used instruments.

Achievement goal theorists (Ames, 1984, 1992; Duda, 2001; Nicholls, 1989; Roberts, 1993; Treasure, 2001) have emphasized that individual goal orientations (e.g., task and ego goal orientations) are distinct from perceptions of the motivational climate. Thus, for example, individual students can have task and ego goal orientations while the climate may place greater emphasis on learning and task involvement or on social comparison, performance, and ego involvement. In distinguishing between individual goal orientations and classroom motivational climates, Ames (1992) suggested that learning goals were reinforced when tasks were diverse, interesting, personally meaningful, challenging, and gave students a sense of control. More generally, there is widespread acceptance that motivational climates created by parents, teachers, coaches—as well as many other characteristics that influence climate—can influence individual motivation and motivational goal orientations. Although goal orientations have been examined in conjunction with many motivational outcomes, limited research has focused on youth involvement in physical exercise. Whereas several studies have reported a positive relationship between task orientation and exercise intentions (e.g., Biddle, Soos & Chatzisarantis, 1999; Papaioannou & Theodorakis, 1996), the present investigation is apparently the first to relate goal orientations to a coherent profile of motivational constructs for a large, nationally representative sample of students from upper-primary, middle and high schools.

Process-related factors and self-conceptions of capacity

Logically connected to goal orientation and one's perceptions of the motivational climate is one's involvement in the process of physical activity. Two aspects of the process that are of interest in this study are task enjoyment and effort expenditure. Research into intrinsic motivation suggests that these two constructs are closely aligned to task motivational orientation (e.g., Lepper, 1988, Marsh, Craven, et al., 2003; Ryan & Deci, 1989) and by implication are deemed to be relevant in the present investigation.

Similarly, one's self-conceptions of physical ability and capacity are hypothesized to be relevant to our integrative approach to the study of motivation. Positive self-concept is valued as a desirable outcome in many disciplines: educational, developmental, clinical and social psychological, and in areas such as sport and health. Self-concept is frequently posited as a mediating variable that facilitates the attainment of other desired outcomes such as physical activity, exercise adherence, or health-related physical fitness. However, early self-concept instruments focused on global self-concept and none provided a clearly interpretable measure of physical self-concept (Marsh, 1997). More recently, there has been a stronger emphasis on physical self-concept measures that are designed specifically for sport and exercise settings (Fox & Corbin, 1989; Marsh, 1997, 2002), providing clear evidence for their convergent and discriminant validity in relation to other self-concept domains (e.g., academic) and sport/exercise outcome measures.

Behavior

Increased levels of physical activity are frequently envisaged outcomes of physical educational classes. Motivational models such as Ajzen's (1988) *theory of planned behavior* have been influential in determining how positive attitudes towards exercise and behavioral intentions are translated into actual exercise behavior. This theory posits that an individual's intention to engage in a given behavior is the most immediate predictor of that behavior. Intentions are largely shaped by attitudes towards behavior, that is, people's assessment of their beliefs regarding the target behavior's effectiveness in producing outcomes and an evaluation of these outcomes. Intention and behavior are also affected by perceived behavioral control, a construct representing people's assessment of their capacities concerning their behavioral engagement. Meta-analytic reviews in exercise behavior indicate medium to large effect sizes for the intention-behavior, attitude-intention and perceived behavioral control-intention relationships (Hagger, Chatzisarantis &

Biddle, 2002). Accordingly, we proposed planned and actual behavior as the final components in our integrative framework. Specifically, we included four constructs derived from this theoretical perspective: positive attitudes towards exercise, perceived control, behavioral intentions and actual behavior in terms of physical exercise (Theodorakis, 1994).

Age, Gender, and Test-retest correlation

Age effects

Developmental researchers (e.g., Harter, 1992) have been concerned about developmental trends in which levels of intrinsic motivation declined during late primary and middle school. Marsh (1989) reported a systematic decline in self-concept during early school years that generalized across academic and non-academic components. Marsh (1990; Marsh, Craven & Debus, 1998; also see Stipek & MacIver, 1989) argued that, whereas young children have extremely high self-concepts, they develop more realistic appraisals of their relative strengths and weaknesses with age, and this added experience is apparently incorporated into their self-concepts. With increasing age children maintain high self-concepts in areas of self-perceived strength but show declines in other areas. Hence, self-concepts in each domain systematically decline with age when averaged across responses by a representative sample of children.

Gottfried, Fleming, and Gottfried (2001) demonstrated that test-retest correlation was moderate for academic intrinsic motivation over the period from middle elementary school through high school, but that there was a systematic largely linear decline in intrinsic motivation over this period. Lepper, Sethi, Dialdin and Drake (1996) found a decline in intrinsic motivation that varied somewhat for different components of this construct and few systematic age-related differences in extrinsic motivation. Lepper et al. also emphasized that this decline in intrinsic motivation seemed to be specific to academic motivation since there was little evidence of a decline in intrinsic motivation for non-school-related activities. Furthermore, they suggested that there was a steady decline in intrinsic motivation with increasing age and year in school, but an additional, relatively larger decline was associated with the transition from elementary to middle school. Greek physical education research shows similar trends in that from the end of elementary school until the end of high school there is a steady decline in students' intrinsic motivation, ability-related beliefs and values, and task orientation (Papaioannou, 1997; Diggelidis & Papaioannou, 1999).

Gender effects

The present study was conducted amongst a large, nationally representative group of young people and therefore is well placed to resolve some of the inconsistencies emerging in relation to gender and motivation. Researchers (e.g., Dweck, 1986; Thorkildsen & Nicholls, 1998) have noted a lack of consistency in research-based reports of gender differences in motivational orientations. A reasonably consistent finding is that boys tend to be more competitively oriented whereas girls are more cooperatively oriented (e.g., Martin, 2001, 2003; Owens & Straton, 1980). Also, males are more likely to attribute academic success to ability whereas girls are more likely to attribute success to effort (e.g., Ames, 1984). Consistent with this pattern of results, Thorkildsen and Nicholls (1998) found significant gender differences for fifth-graders such that boys scored higher in ego-orientation, alienation, and beliefs in extrinsic causes of success whereas girls scored higher in task-orientation and beliefs in interest and effort as the causes of success. Placing their findings in a broader context, however, Thorkildsen and Nicholls indicated that they had not found gender differences in motivational orientations in studies conducted with younger (2nd grader) or older (adolescent) students. Owens and Straton (1980), however, indicated that gender differences in Cooperative, Individual, and Competitive Orientations did not interact with year in school for a large group of students in grades 4-11. In contrast, Ablard and Lipschultz (1998) reported that for a group of high achieving, 7th grade students, girls had significantly higher Learning orientations than boys, whereas there were no significant differences for Performance orientations.

In the physical activity context, different studies have suggested that boys have higher ego-orientation and lower task-orientation (Duda, 1989; Lintunen, Valkonen, Leskinen & Biddle, 1999), girls have higher task-orientation but no gender differences in ego-orientation (Newton & Duda, 1993), higher ego-orientation for males but no gender differences in task-orientation (Digelidis & Papaioannou, 1999) and no gender difference in either task- or ego-orientation (Ebbeck & Becker, 1994). It should be noted though, that whenever statistically significant gender differences emerged, the magnitude was never large. Hence, as noted by others, the research literature does not provide a clear picture about gender differences in motivational orientations and their development. On the other hand, research across different cultures has consistently shown that males have higher perceptions of sport competence than females (Digelidis & Papaioannou, 1999; Lintunen et al., 1999; Marsh, Hey, Roche & Perry, 1997).

Test-retest correlation

Existing research implies that goal orientations are not fixed over time. For example, Williams (1998) assessed athletes' goal orientations early and late in a competitive season and reported test-retest correlation coefficients of .44 and .64 for task and ego orientation respectively. Duda (2001) reviewed research showing higher test-retest correlations for a 3-week period, but noted that the values were lower when the test-retest period was one year. She suggested that goal orientations were not personality traits but rather dispositional tendencies that were malleable and were impacted by situational factors. We see the inclusion of a diversity of motivation-related constructs in the present investigation as an ideal opportunity to conduct a broader study of test-retest correlation than that previously conducted. In particular, we apply sophisticated methods that take into consideration some of the more complex issues related to the issue of test-retest correlation (e.g., correlated uniquenesses) and in so doing are able to gauge more accurately the longitudinal profile of motivation in the physical activity setting. We note that test-retest stability over relatively short periods of time based on responses by the same participants on multiple occasions is a separate issue from evaluation of age effects based on cross-sectional data covering a wide range of preadolescent and adolescent ages as considered in the present investigation.

Method

Procedures and Sample

Participants were 2786 students (50% males) from 200 physical education classes at different levels of schooling (29% primary school, 36% middle school, and 35% high school). The schools were randomly selected from the total number of schools from nine different geographical areas of Greece, involving both urban and suburban areas and different social classes, in order to be nationally representative of Greek students. T1 variables were collected shortly after the start of the school year (September - October 1998) whereas T2 variables were collected near the end of the school year (April - May 1999). T1 was at least 5 weeks after the beginning of the school year so that most students had at least 10 class sessions with the same teacher. At both times the anonymous questionnaires were distributed by nine research assistants and were completed in the students' classes. Student consent and permission from the Ministry of Education and the school authorities were required. An important complication in the present investigation was the requirement (by law of the Greek Ministry of Education) that all questionnaires should be completed anonymously. Hence, for purposes of the present investigation, T1 and T2 cases were matched on the basis of class identification, gender, and date of birth. Because not all students provided a proper date of birth on both occasions, many cases could not be matched. For present purposes, we only considered classes for which there were at least 10 students at T1, at least 10 students at T2, and at least five successfully matched cases with data for T1 and T2. Excluded were participants who did not participate in both data collections (a few schools had data from only T1) and students who did not have the same physical education teacher at T1 and T2.

Measures

For present purposes we consider 11 factors measured at Time 1 (T1) and again at Time 2 (T2): two climate factors—task-involving and ego-involving, two motivational goal orientation factors, two intrinsic motivation measures, four measures of planned exercise behavior, one measure of physical self-concept, and three background variable factors. Hence, there were a total of 25 factors (11 T1 factors, 11 T2 factors, 3 background variable factors) inferred on the basis of responses to 95 items (see Appendix).

Task- and ego-involving climates. Students were asked to evaluate their perceptions of the motivational climate in their physical education class. The 7-item Task-involving climate scale (Papaioannou, 1994; see Appendix) had internal consistency reliability estimates of .77 (T1) and .81 (T2). The 6-item Ego-involving climate scale (Papaioannou, 1998; see Appendix) had internal consistency reliability estimates of .83 (T1) and .85 (T2).

Task and Ego goal orientations. The 7-item task orientation in physical education classes (Duda & Nicholls, 1992; adapted for Greek physical education by Papaioannou & Macdonald, 1993; see Appendix) had internal consistency reliability estimates of .83 (T1) and .88 (T2). The 6-item ego orientation in physical education classes (Duda & Nicholls, 1992; adapted for Greek physical education by Papaioannou & Macdonald, 1993; see Appendix) had internal consistency reliability estimates of .78 (T1) and .81 (T2).

Intrinsic motivation. The 3-item scale, Enjoyment in physical education classes (McAuley, Duncan & Tammen, 1989; adapted for Greek physical education by Diggelidis & Papaioannou, 1999; see Appendix), had internal consistency reliability estimates of .84 (T1) and .89 (T2). The 3-item scale, Effort in physical education classes (McAuley, Duncan & Tammen, 1989; adapted for Greek physical education by Diggelidis & Papaioannou, 1999; see Appendix), had internal consistency reliability estimates of .83 (T1) and .86 (T2).

Planned Exercise Behavior. The 3-item positive Attitudes towards exercise (Theodorakis, 1994; see Appendix) had internal consistency reliability estimates of .58 (T1) and .75 (T2). For this one scale, it appears that the reliability of particularly the T1 measure has not reached an acceptable level suggesting that interpretations should be made cautiously and that future research with this construct should consider revision of the scale. The 3-item Intentions to exercise (Theodorakis, 1994; see Appendix) had internal consistency reliability estimates of .81 (T1) and .87 (T2). The 3-item Perceived behavioral control towards exercise (Theodorakis, 1994; see Appendix) had internal consistency reliability estimates of .83 (T1) and .90 (T2). Because the scale for Actual Exercise Behavior in the last month was based on responses to a single item (Theodorakis, 1994; see Appendix), it was not possible to estimate reliability.

Self-concept. The 5-item physical self-concept (5 items) (Fox & Corbin, 1989; adapted in Greek by Diggelidis & Papaioannou, 1999; see Appendix) had internal consistency reliability estimates of .80 (T1) and .82 (T2).

Statistical Analysis

Confirmatory factor analysis. Confirmatory factor analysis models were conducted with LISREL 8 (version 8.54) using maximum likelihood estimation. A detailed presentation of the conduct of CFA and evaluation of goodness of fit is beyond the scope of the present investigation and is available elsewhere (e.g., Joreskog & Sorbom, 1993; Marsh; in press; Balla, & Hau, 1996). In evaluating goodness of fit of alternative models, we emphasized the root mean square error of approximation (RMSEA), but also present the Tucker-Lewis index (TLI), the relative noncentrality index (RNI), the χ^2 test statistic, and an evaluation of parameter estimates. For RMSEAs, values of less than .05 and .08 are taken to reflect a close fit and a reasonable fit, respectively (see Joreskog & Sorbom, 1993; Marsh et al., 1996). The TLI and RNI vary along a 0-to-1 continuum in which values greater than .90 and .95 are typically taken to reflect acceptable and excellent fit to the data.

In preliminary analyses we evaluated the factor structure of the 11 factors collected at T1 (see Appendix) separately for upper-primary, middle, and high school students. The fit of the overall model was good for each group considered separately (see Table 1): primary school students (TLI = .958, RNI = .962, RMSEA = .035), middle school students (TLI = .968, RNI = .971, RMSEA = .032), and high school students (TLI = .972, RNI = .974, RMSEA = .033). Furthermore, multiple group analyses provided good support for the invariance of factor loadings across the three groups (RMSEA = .034 for models with no invariance constraints and .034 for models with factor loadings constrained to be equal across the three groups).

The main analyses were conducted on the combined set of T1 and T2 responses, allowing us to evaluate the test-retest correlation. All these analyses included correlated uniquenesses posited a priori to account for method effects associated with the same items administered on different occasions in longitudinal research. Correlated uniquenesses resulted when there was unique variance associated with responses to one measured variable that was related to responses to another measured variable that could not otherwise be explained by the proposed factor structure. Marsh and Hau (1996; also see earlier discussion by Joreskog, 1979) emphasize that if the same measurements are used on multiple occasions, as is typical in longitudinal research, the corresponding residual error variables will tend to be correlated and, in order to get accurate estimates of relations among the constructs, correlations among errors must be included in the model and should constitute the a priori model. In preliminary analyses, the inclusion of these correlated uniquenesses was supported by modestly better fits to the data and, in particular, their exclusion would have positively biased the corresponding correlation estimates. Their inclusion, however, had no substantively important effect on the pattern of parameter estimates, suggesting that the inclusion of correlated uniquenesses in this study was not a critical issue. In order to facilitate the substantive import of the results, only the models with correlated uniquenesses were presented. It is important to emphasize that all correlated uniquenesses were for the same item administered at T1 and T2, were hypothesized a priori, and did not include correlated uniquenesses between different (non-matching) items.

Data Transformations and Interaction Effects. Several data transformations were conducted to facilitate interpretations and infer interaction effects. Because there were moderate amounts of nonnormality in many of the variables, we began by using a normalizing transformation (SPSS, 1999) on each of the variables and then standardized (z-scoring) all variables to have M = 0, SD = 1 across the entire sample. Also, even though many of the variables reflected responses to Likert response scales, the use of polychoric correlations is typically not an appropriate option unless a weighted least square estimation is used instead of the typical maximum likelihood estimation. However, such estimation procedures are typically only useful with extremely large Ns. Thus, for example, these procedures may require sample sizes as large as 5,000 cases (e.g., Bentler & Dudgeon, 1996; Curran, West, & Finch, 1996; Marsh, Wen & Hau, 2004). Hence, even the substantial sample size for the total sample in the present investigation was not really sufficient—particularly for analyses based on multiple group analyses had sample sizes that were much smaller than this value. Hence, the use of normalization seemed a reasonable compromise. Again, however, it is important to emphasize that the parameter estimates that are the major focus of the present investigation are extremely robust in relation to non-normality. Product terms were used to test interaction effects. In constructing these interaction effects, we used the product of individual (z-score) standardized variables, and the product terms were not re-standardized.

For large-scale studies, the inevitable missing data is a potentially important problem, particularly when the amount of missing data exceeds 5% (e.g., Graham & Hoffer, 2000). In the present investigation, for students with successfully matched T1 and T2 responses, there was little missing data (less than 1% at both T1 and T2). Nevertheless, in the methodological literature on missing data (e.g., Brown, 1994; Graham & Hoffer, 2000; Little & Rubin, 1987), there is a growing consensus that the imputation of missing observations has important advantages over traditional approaches such as pairwise, listwise deletion for missing data, leading us to implement the Expectation Maximization (EM) Algorithm, the most widely recommended approach to imputation for missing data, as operationalized using missing value analysis in SPSS (SPSS, 1999). In the present investigation, however, we explored a variety of alternative approaches to this problem which showed that results based on the EM algorithm that we used were very similar to those based on pairwise and listwise deletion methods for missing data—as would be expected to be the case when there was so little missing data.

MIMIC approach to evaluation of gender and age effects. In the present investigation, we used the MIMIC approach to CFA to evaluate gender and age differences in latent constructs representing 11 motivation factors assessed at T1 and again at T2. As described by Kaplan (2000; also see Jöreskog & Sörbom, 1996; Marsh, Ellis, Parada, Richards & Heubeck, 2005), the MIMIC model approach is like a multivariate regression model in which latent variables (e.g., multiple motivation factors) are "caused" by discrete contrast or grouping variables (e.g., age, gender, age x gender) that are each represented by a single indicator. This approach is clearly stronger than a traditional MANOVA approach that is based on measured variables (i.e., scale scores) that are assumed to be measured without error rather than latent variables arising from the CFA.

The MIMIC CFA approach is also much more flexible than the traditional multivariate analysis of variance (MANOVA) approach in allowing a mixture of continuous and discrete independent variables (i.e., contrast, background or grouping variables) and their interactions. Although like a multiple regression approach to ANOVA or MANOVA, the MIMIC approach has the important advantage in that the dependent variables are latent variables based on multiple indicators with appropriate control for measurement error. The MIMIC approach also leads more naturally to recent developments in CFA models such as latent growth modeling (for further discussion, see Kaplan, 2000). The MIMIC approach is particularly advantageous in the present investigation in that it allows us to incorporate tests of the factor structure, test-retest correlations, age effects and gender effects into a single analytic framework rather than introducing new, potentially suboptimal statistical analyses (e.g., traditional MANOVAs of motivation scale scores to evaluate gender and age effects) to address different research questions.

For the MIMIC approach as well as related correlational approaches, it is important to distinguish between statistical significance and practical importance. The focus of our statistical analyses is on statistical significance and we avoid use of the term "significant" in relation to practical importance. It is difficult, however, to find well-accepted criteria of how to evaluate the size of effects. For example, recognizing the unavoidable subjectivity of such interpretations, Glass, McGaw, and Smith (1981) stated that "there is no wisdom whatsoever in attempting to associate regions of the effect-size metric with descriptive adjectives such as 'small,' 'moderate,' or 'large' and the like" (p. 104).

Results & Discussion

Confirmatory Factor Analysis and Correlations Among Variables

We began with a brief evaluation of the factor structure underlying our constructs and the pattern of relations among these latent factors. For present purposes, we considered 11 factors measured at T1 and again at T2 (i.e., 22 latent factors). For purposes of this study, we discuss these 11 factors as two climate factors (individual perceptions of Task-involving and Ego-involving climates), two motivational orientation factors (Task and Ego), and 7 other outcome factors. All but one of these factors (Behavior) were multi-item factors in which the latent factor was inferred on the basis of multiple indicators. In addition, there were three background variable factors (age, gender, and age x gender interaction) that were single-indicator factors. Hence, the overall factor structure consisted of 95 measured variables that were used to infer 25 (11 T1 factors, 11 T2 factors, and 3 background variable factors) latent factors. We fit a highly restrictive a priori structure in which each indicator was allowed to load only on the a priori factor that it was designed to measure. Results (Table 2) showed that the factor structure was well defined, as all factor loadings were highly statistically significant and substantial. Moreover, the correlations among factors (Table 3) formed a logical pattern of relations, and the goodness of fit was very good in relation to traditional guidelines (e.g., RMSEA = .034). Of particular interest were factor correlations among background variables and the outcome variables collected at T1 and T2 (Table 3).

Age was statistically significantly negatively related to nearly all the outcome variables and the individual class climate factors at both T1 and T2. The correlations were particularly negative for the individual perceptions of task-involving climate, enjoyment, and effort (-.35 to -.48 at T1 and T2, Table 3) but were also significantly negative for task orientation, exercise intentions, perceived behavioral control, actual behavior, and physical self-concept. The only factors that were not negatively related to age at both T1 and T2 were individual perceptions of Ego-involving Climate, Ego Orientation, and Attitudes Towards Exercise (although even these factors were significantly negatively related to age at either T1 or T2). In summary, none of the outcome measures were positively correlated with age and most were substantially negatively correlated with age.

Boys had systematically more positive outcomes for most of the variables considered here (Table 3 where negative relations with gender, 1=boys, 2=girls, reflect higher scores for boys). Whereas most of these gender differences were small in size, the largest correlations were for Physical Self-concept (-.32 and -.31 at T1 and T2 respectively). Whereas boys had somewhat stronger ego orientations and perceived the classroom climate to be somewhat more ego-involving than girls, there were no statistically significant gender differences for task orientation and individual student perceptions of task-involving climates. In contrast to other outcomes, girls had slightly more positive attitudes

Table 1. Summary of Goodness of Fit For Confirmatory Factor Analysis Models

Model	$\chi 2$	DF	TLI	RNI	RMSEA	Description
Total Group Analyses						
1	17342.777	4031	.978	.976	.034	Total Group
Separate Group Analyses						
2a	7744.072	3891	.958	.962	.035	Primary School
2b	8276.450	3891	.968	.971	.032	Middle School
2c	7795.199	3891	.972	.974	.033	High School
Multiple Group Invariance Tests						
3a	23815.721	11673	.967	.970	.034	No Invariance
3b	24230.441	11813	.967	.969	.034	Model 3a with factor loadings invariant
3c	24853.906	11859	.965	.968	.034	Model 3b factor variances invariant
3d	26481.334	12365	.964	.965	.035	Model 3c factor covariances invariant
3e	28138.006	12633	.960	.961	.036	Model 3d with uniqueness invariant (total invariance)

Note: RNI = relative noncentrality index, TLI = Tucker-Lewis index, RMSEA = root mean square error of approximation. Total N = 2786 (811 elementary, 1076 junior high, 899 high school) students. Model 1 was based on 95 items used to infer 25 latent constructs. In separate models students in each age group, age and its interaction with gender (both included as single-item factors in Model 1) were excluded so that there were only 93 items used to infer 23 latent constructs. In the final set of models (3a – 3e), the results provide strong support for the invariance of factor loadings across the three age groups (the comparison of Models 3a and 3b) and reasonable support for the complete invariance of all parameter estimates.

Table 2. Confirmatory Factor Analysis Results: Factor Loadings[a]

Climate		Orientation		Intrinsic Motivation		Planned Behavior				Self-concept	Background/Demographic		
TSK	EGO	TSK	EGO	ENJ	EFF	EAT	EIN	PBC	BEH	PSC	Age	Sex	AxS
Time 1 Constructs													
CTSK1 .61	CEGO1 .39	OTSK1 .63	OEGO1 .56	ENJ1 .78	EFF1 .81	EAT1 .63	EIN1 .81	PBC1 .70	BEH1 1.0	PSC 1 .62			
CTSK2 .66	CEGO2 .57	OTSK2 .38	OEGO2 .74	ENJ2 .82	EFF2 .74	EAT2 .47	EIN2 .85	PBC2 .80		PSC 2 .74			
CTSK3 .64	CEGO3 .76	OTSK3 .62	OEGO3 .79	ENJ3 .81	EFF3 .81	EAT3 .57		PBC3 .84		PSC 3 .68			
CTSK4 .65	CEGO4 .72	OTSK4 .60	OEGO4 .81							PSC 4 .73			
CTSK5 .64	CEGO5 .62	OTSK5 .69	OEGO5 .58							PSC 5 .57			
CTSK6 .68	CEGO6 .46	OTSK6 .48	OEGO6 .56										
CTSK7 .69		OTSK7 .57											
Time 2 Constructs													
CTSK1 .69	CEGO1 .49	OTSK1 .67	OEGO1 .61	ENJ1 .85	EFF1 .84	EAT1 .72	EIN1 .86	PBC1 .81	BEH1 1.0	PSC 1 .64	Age 1.0	Sex 1.0	AxS 1.0
CTSK2 .77	CEGO2 .66	OTSK2 .47	OEGO2 .74	ENJ2 .88	EFF2 .77	EAT2 .67	EIN2 .89	PBC2 .86		PSC 2 .73			
CTSK3 .70	CEGO3 .80	OTSK3 .65	OEGO3 .82	ENJ3 .84	EFF3 .86	EAT3 .72		PBC3 .91		PSC 3 .71			
CTSK4 .72	CEGO4 .76	OTSK4 .65	OEGO4 .84							PSC 4 .75			
CTSK5 .68	CEGO5 .69	OTSK5 .73	OEGO5 .59							PSC 5 .63			
CTSK6 .71	CEGO6 .45	OTSK6 .57	OEGO6 .56										
CTSK7 .71		OTSK7 .62											

Note: TSK-C = Task-involving Climate ; EGO-O = Ego-involvinnnng Climate ; TSK-O = Task Goal Orientation ; EGO-O = Ego Goal Orientation ; ENJ = Enjoyment ; EFF = Effort ; EAT = Exercise Attitudes; EIN = Exercise Intent; PBC = Perceived Behavioral Control ; BEH = Behavior ; PSC = Physical Self-concept; AxS = age-by-gender interaction. All parameter estimates are presented in completely standardized form. The confirmatory factor analysis, based on responses to 95 items that were posited a priori to measure 25 factors, provided a good fit to the data (see Model 1, Table 1).

[a] Each factor was based on responses between 1 and 7 items. Hence, a factor with 7 items will have 7 factor loadings, whereas a scale with only three items will have only three factor loadings (and blanks for the remaining spaces). Five of the 25 factors (behavior at Time 1 and Time 2, age, gender, and the age-by-gender interaction) were based on a single item and, as a consequence, have standardized factor loadings of 1.0. In expanded form, this factor loading matrix will be a 95 (items) x 25 (factor) matrix, but is presented here in condensed form to conserve space.

Table 3. Confirmatory Factor Analysis Results: Factor Correlations

	Time 1 Factors											Time 2 Factors											Background		
	Climate		Orient									Climate		Orient											
	TSK	EGO	TSK	EGO	ENJ	EFF	EAT	EIN	PBC	BEH	PSC	TSK	EGO	TSK	EGO	ENJ	EFF	EAT	EIN	PBC	BEH	PSC	Age	SEX	AxS
Climate Perceptions																									
TSK-C	1.00																								
EGO-C	.01	1.00																							
Motivation Goal Orientations																									
TSK-O	.63	.08	1.00																						
EGO-O	.04	.34	.14	1.00																					
Intrinsic Motivation																									
ENJ	.69	.05	.59	.03	1.00																				
EFF	.62	.16	.66	.10	.71	1.00																			
Planned Behavior																									
EAT	.25	-.06	.39	.04	.27	.24	1.00																		
EIN	.34	.04	.46	.05	.40	.41	.51	1.00																	
PBC	.37	.07	.44	.07	.45	.46	.43	.92	1.00																
BEH	.15	.04	.24	.12	.21	.23	.20	.42	.42	1.00															
Self-concept																									
PSC	.38	.40	.39	.35	.40	.46	.19	.37	.40	.28	1.00														
Climate Perceptions																									
TSK-C	.57	-.00	.35	.02	.47	.42	.10	.24	.29	.13	.25	1.00													
EGO-C	-.08	.41	-.05	.19	-.01	.01	-.05	.01	.03	.04	.20	-.01	1.00												
Motivation Goal Orientations																									
TSK-O	.43	.02	.53	.07	.43	.50	.23	.36	.36	.20	.32	.58	.03	1.00											
EGO-O	-.03	.24	.00	.47	-.02	-.00	.02	.02	.03	.09	.25	.04	.45	.13	1.00										
Intrinsic Motivation																									
ENJ	.50	.03	.37	.01	.58	.47	.12	.29	.33	.14	.28	.69	-.02	.56	.01	1.00									
EFF	.45	.13	.42	.07	.48	.62	.11	.29	.33	.15	.35	.60	.13	.64	.11	.71	1.00								

	Time 1 Factors											Time 2 Factors											Background		
	Climate		Orient									Climate		Orient											
	TSK	EGO	TSK	EGO	ENJ	EFF	EAT	EIN	PBC	BEH	PSC	TSK	EGO	TSK	EGO	ENJ	EFF	EAT	EIN	PBC	BEH	PSC	Age	SEX	AxS
Planned Behavior																									
EAT	.18	-.11	.23	-.00	.15	.15	.37	.22	.19	.13	.05	.19	-.09	.32	-.02	.21	.17	1.00							
EIN	.30	.04	.35	.05	.35	.37	.27	.58	.59	.34	.35	.34	.02	.47	.07	.39	.42	.47	1.00						
PBC	.33	.08	.35	.07	.38	.42	.23	.57	.64	.35	.40	.39	.05	.49	.07	.42	.46	.38	.93	1.00					
BEH	.19	.09	.27	.08	.26	.30	.14	.38	.38	.41	.32	.23	.06	.32	.10	.26	.30	.18	.52	.52	1.00				
Self-concept																									
PSC	.19	.12	.24	.10	.25	.26	.12	.25	.27	.19	.58	.32	.41	.39	.43	.34	.40	.08	.40	.42	.35	1.00			
Background																									
Age	-.35	-.15	-.26	-.06	-.40	-.48	-.03	-.25	-.34	-.08	-.32	-.36	-.02	-.30	-.01	-.38	-.45	-.08	-.32	-.39	-.27	-.25	1.00		
SEX[a]	-.02	-.13	.01	-.08	-.09	-.08	.03	-.09	-.10	-.15	-.30	-.04	-.15	-.01	-.15	-.08	-.10	.09	-.09	-.10	-.18	–.31	.08	1.00	
AxS	-.00	.01	.04	.01	-.07	-.05	.03	-.00	-.02	-.04	-.08	.00	.01	-.01	-.00	-.05	-.08	.00	-.08	-.08	-.07	-.04	.03	-.03	1.00

Note: TSK-C = Task-involving Climate ; EGO-O = Ego-involving Climate ; TSK-O = Task Goal Orientation ; EGO-O = Ego Goal Orientation ; ENJ = Enjoyment ; EFF = Effort ; EAT = Exercise Attitudes; EIN = Exercise Intent ; PBC = Perceived Behavioral Control ; BEH = Behavior ; PSC = Physical Self-concept; AxS = age-by-gender interaction. All parameter estimates are presented in completely standardized form. All parameter estimates greater than .05 in absolute value are statistically significant ($p < .05$, two-tailed). Test-retest correlations appear in bold.

a The 24 effect sizes (*d*) corresponding to the 24 correlations between sex and the remaining 24 variables (in the order presented in this table) respectively are:
-.04 -.26 .02 -.16 -.18 -.16 .06 -.18 -.20 -.30 -.63 -.08 -.30 -.02 -.30 -.16 -.20 .18 -.18 -.20 -.37 -.65.

towards exercise than boys (although this gender difference was only significant at T2). The age x gender interactions (represented by the age x gender crossproduct term) were mostly small, although all of the statistically significant effects indicated that gender differences in favor of boys increased with age.

Perceptions of Task-involving and Ego-involving Climates were not significantly correlated at either T1 or T2 (Table 3), whereas Task and Ego goal orientations were somewhat positively correlated (.14 at T1 and .13 at T2). Perceptions of Task-involving climate were substantially correlated with Task Orientation (.63 at T1, .58 at T2) but were nearly uncorrelated with Ego Orientation (.04 at T1 and T2). Similarly, individual perceptions of Ego-involving climate were substantially correlated with Ego Orientation (.34 at T1, .45 at T2) but were nearly uncorrelated with Task Orientation (.08 at T1, .03 at T2).

Perceptions of Task-involving Climate were statistically significantly correlated with all the remaining outcome variables at both T1 and T2 (Table 3). The largest correlations were with the two intrinsic motivation factors (.60 to .69 at T1 and T2). It was interesting to note that the perceptions of Task-involving Climate were as highly correlated with the two intrinsic motivation factors—Effort and Enjoyment—as with Task Orientation. Inspection of the wording of the items, however, revealed that students were actually making ratings of the physical education class (e.g., the physical education lesson is fun) rather than their own intrinsic motivation for pursuing exercise and physical activity, whereas ratings of the Task Orientation items were not made in relation to students' physical education classes even though the content of these items appeared to be more closely related to the Task-involving Climate items. Although many of the correlations between individual perceptions of Ego-involving Climate and the other outcomes were close to zero, most were statistically significantly positive at T1, T2 or both T1 and T2. Nevertheless, the magnitude of these relationships was small. Only Attitudes Toward Exercise was significantly negatively correlated with individual perceptions of Ego-involving Climate (-.08 and -.11 at T1 and T2 respectively). Whereas Task-involving Climate was more highly correlated with most of the other outcome variables than Ego-involving Climate, both the climate factors were substantially positively correlated with Physical Self-concept.

At both T1 and T2, there was a systematic and predictable pattern of relations among the different outcome variables. As already noted, the correlations were nonsignificant between the two climate variables and small positive correlations were found between the two motivational goal orientation variables. The remaining outcome variables were all positively correlated. The highest correlations were between Exercise Intentions and Perceptions of Control of Exercise Behavior (.92 and .93 at T1 and T2 respectively), and these variables were each positively correlated with actual Exercise Behavior (.42 to .52). The two intrinsic motivation factors, Enjoyment and Effort, were also highly correlated (.71 at both T1 and T2).

Because the two classroom climate factors and the other nine outcome factors were both measured at T1 and T2, the results also included test-retest correlations. These test-retest correlation coefficients varied from a low of .37 (Attitudes towards Exercise) to .64 (Perceived Behavioral Control). Because these correlations were based upon latent factors that were corrected for measurement error, the correlations reflect test-retest correlation that was unconfounded with unreliability. Although statistically significant and substantial, none of these test-retest correlation coefficients approach 1.0.

Summary and Implications

In summary, the results of this confirmatory factor analysis supported the construct validity of the constructs that form the basis of the present investigation. The excellent fit of the a priori confirmatory factor analysis model and the factor loadings indicated that each of the constructs considered here was well defined. Further support for construct validity came from the systematic and predictable pattern of relations among the different factors. In considering these results in further detail, we highlight what we considered to be some of the key findings.

It is important to recognize that motivation in the physical activity domain is negatively related to age and that this seems to parallel declines in academic motivation (Gottfried et al., 2001). Moreover, this decline seems to interact with gender such that girls' physical motivation declined more markedly than boys' motivation. This, coupled with the fact that girls' physical activity also declined more markedly than boys' activity, suggested that gender and the student's developmental level were relevant to considerations of physical activity and underlying physical motivation. These findings were similar to reports in the USA (McKenzie, 2001), but they should not be interpreted to imply that girls are not interested in physical activity (Corbin, 2002). Rather, there is a need to rethink the sport-oriented curriculum of current physical education lessons and the competitive structure of youth physical activity contexts, which are preoccupied with activities that are either perceived to be masculine in nature or uninteresting to young females (Corbin, 2002).

If, as we argue (Papaioannou, Marsh, & Theodorakis, 2004), students' motivational orientations and perceptions of the climate predict their level of activity, then these are feasible points of intervention. In view of the findings in the present study, this brings into consideration strategies aimed at enhancing students' task focus and promoting a task-oriented climate (Duda, 1996). Enhancing students' task focus can be achieved through promoting a focus on process more than outcomes, personal bests—competing with one's previous performance more than competing with others, and showing students that the quality and quantity of effort (rather than ability) are the key means of improvement and accomplishment (Craven, Marsh, & Debus, 1991; Martin, Marsh, & Debus, 2001, 2003). In terms of climate, effective strategies include promoting a classroom climate of cooperation, self-improvement, and personal bests (Qin, Johnson, & Johnson, 1995). Further details for the promotion of a task-involving climate in physical education can be found elsewhere (Papaioannou & Goudas, 1999; Treasure & Roberts, 1995).

The present findings imply that policies aimed at reducing the decline of task-involving climate should be a priority in Greek physical education. This would be likely to reduce the decline of adolescents' interest in physical activities with age. Despite the absence of large studies assessing the evolution of motivational climate in physical education in other countries, the available evidence (Gottfried et al., 2001; McKenzie, 2001) suggests that our conclusions are generalizable across most western societies.

Nevertheless, there are a number of potential limitations to the present investigation (reliance on a Greek sample, potential problems with missing data, the small number of items used to measure some of the constructs) that warrant caution in generalizations based on our results and the need for further research. Furthermore, although there is insufficient research on how to eliminate the decline of task-involving climate with age, theory predicts that both task and teaching practices contribute to the formation of motivational climate (Ames, 1992). Nevertheless, it seems doubtful whether an intervention in teaching practices without changing the masculine or competitive nature of physical activities would increase the motivation of girls or of students perceived to be of low physical ability, respectively.

It is interesting that individual and climate constructs were similarly correlated with process, behavior, and self-conceptions. In the first instance, this shows that climate was an important factor in students' enjoyment, effort, and behavior. We recognize that there is increasing research into climate-related issues but emphasize that it is an important construct to pursue in future research, and that it attests to the importance of intervention at individual and climate levels. Notwithstanding this, it is important to interpret these findings in the light of the fact that students' ratings of climate were obviously based on their own perceptions of climate, so that the potential confounding of individual orientation and individual perception of climate is an issue that requires further investigation (Papaioannou, Marsh, & Theodorakis, 2004).

Importantly, whereas the results of the present investigation are clearly supportive of the positive effects of task orientations and task-involving climates, there is no evidence that ego orientations and ego-involving climates are bad. Indeed, relations between these constructs and other outcomes in the present investigation tended to be slightly positive. Hence, whereas relations involving these ego constructs are less positive than those involving the corresponding task constructs, there is no evidence that these ego constructs have negative effects. It is also relevant to emphasize that it is possible that other motivational orientations not considered in the present investigation (e.g., social affiliation) are important motivators of physical activity. Undoubtedly, further research is needed on this topic.

In addition to addressing substantive issues related to motivation and associated gender and age effects, and test-retest correlation, the study also advanced the field from a measurement perspective. The study adopted an integrative framework aimed at drawing together a coherent profile of motivation and motivation-related constructs. In our discussion of construct validation at the outset of the paper we noted that robust measures should be psychometrically sound not only when considered separately, but also in the context of other measures. This study was important because it demonstrated that the motivation constructs under analysis were robust even when considered in the context of an integrative framework that comprises individual- and climate-level constructs. Hence, we have identified a strong package of measures that researchers can use with confidence.

Conclusion

In the present study we applied a rigorous construct validation approach to a coherent profile of motivational measures in a physical education setting within the context of a nationally representative longitudinal sample of students covering a wide range of ages. The data have identified a strong factor structure in the context of an integrative framework, noteworthy relationships amongst key constructs with gender and age, clarity regarding the test-retest correlation of these motivation constructs, and a package of sound measures that offers researchers a strong basis upon which to pursue motivation research at individual and climate levels. Taken together, the high quality sampling combined with the rigorous analytical approach and the cohesive package of measures underscore the importance of the study in contributing to current understanding of motivation in the physical education context.

Psychometric rigor is an important criterion for the evaluation of motivation factors in the present investigation. Whereas reliability estimates for the scales would clearly be higher if the length of the scales had been substantially increased, this would have undermined their usefulness. Hence, in relation to its intended purposes, there is a reasonable overall balance between brevity and psychometric rigor. Nevertheless, it is important to note that some of the particularly brief scales did not reach fully acceptable levels of reliability so that further revision would be warranted in their application in further research. Particularly in relation to omnibus measures typically used in sport psychology, the factor structure underlying responses to this instrument is extremely well defined, as evidenced by the excellent fit statistics resulting from the CFAs reported earlier. To the best of our knowledge, there is no other broadly based sport psychology instrument that has such a well-defined factor structure as rigorously assessed with CFA approaches like those used here. Whereas the methodology for testing invariance of factor structures across different groups of respondents is well developed, the instrument considered here is also unique among broadly based sport psychology measures in terms of its demonstrated invariance of factor structure across gender and such a wide range of ages.

Although there is a clear theoretical rationale behind each of the constructs selected for inclusion in the present investigation, the set of scales is theoretically eclectic, cutting across a range of different theoretical perspectives that are important in sport psychology. This is a particular strength in that many constructs in sport psychology are typically evaluated within a narrow theoretical framework, often by the same group of researchers. In this respect, there is greater breadth in the constructs considered here than is typical in sport psychology research. As emphasized by Marsh (1994) in his discussion of the jingle-jangle effects, similar constructs are sometimes given different labels that imply a distinctiveness that may not be warranted, whereas apparently different constructs are sometimes given similar labels that imply a convergence that may also be unwarranted. Clearly, we are not claiming that the set of constructs considered here represent all possible—or necessarily even the most important—constructs in sport psychology. However, the set of constructs provides a useful framework against which to evaluate the convergent and discriminant validity of new and existing sport psychology constructs.

References

Ablard, K. E., & Lipschultz, R. E. (1998). Self-regulated learning in high-achieving students: Relations to advanced reasoning, achievement goals, and gender. *Journal of Educational Psychology, 90,* 94-101.

Ajzen, I. (1988) *Attitudes, personality, and behaviour.* Chicago: Dorsey.

Ames, C. (1984). Achievement attributions and self-instructions under competitive and individualistic goal structures. *Journal of Educational Psychology, 76,* 478-487.

Ames, C. (1992). Classroom goals, structures, and student motivation. *Journal of Educational Psychology, 84,* 261-271.

Bentler, P. M., & Dudgeon, P. (1996). Covariance structure analysis: Statistical practice, theory, and directions. *Annual Review of Psychology, 47,* 563-592.

Biddle, S., Soos, I., & Chatzisarantis, N. (1999). Predicting physical activity intentions using goal perspectives and self-determination theory approaches. *European Psychologist, 4,* 83-89.

Brown, R. L. (1994). Efficiency of the indirect approach for estimating structural equation model with missing data: A comparison of five methods. *Structural Equation Modeling, 1,* 287-316.

Corbin, C. B. (2002). Physical activity for everyone: What every physical educator should know about promoting lifelong physical activity. *Journal of Teaching in Physical Education, 21,* 128-144.

Craven, R. G., Marsh, H. W., & Debus, R. L. (1991). Effects of internally focused feedback and attributional feedback on the enhancement of academic self-concept. *Journal of Educational Psychology, 83,* 17-26.

Curran, P. J., West, S. G., Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1,* 16-29.

Diggelidis, N. & Papaioannou, A. (1999). Age-group differences in intrinsic motivation, goal orientations and perceptions of athletic competence, physical appearance and motivational climate in Greek physical education. *Scandinavian Journal of Medicine & Science in Sports, 9,* 375-389.

Duda, J. L. (1989). Relationship between task and ego orientation and the perceived purpose of sport among high school athletes. *Journal of Sport & Exercise Psychology, 11,* 318-335.

Duda, J. L. (1996). Maximizing motivation in sport and physical education among children and adolescents: the case for greater task-involvement. *Quest, 48,* 290-302.

Duda, J. L. (1998). *Advances in Sport and exercise psychology measurement.* Morgantown, WV: Fitness Information Technology.

Duda, J. L. (2001). Achievement goal research in sport: Pushing the boundaries and clarifying some misunderstandings. In G. C. Roberts (Ed.), *Advances in motivation in sport and exercise* (pp. 129-182). Champaign, IL, Human Kinetics.

Duda, J. L., & Nicholls, J. G. (1992). Dimensions of achievement motivation in schoolwork and sport. *Journal of Educational Psychology, 84,* 290-299.

Dweck, C. S. (1986). Motivation processes affecting learning. *American Psychologist, 41,* 1040-1048.

Ebbeck, V., & Becker, S. (1994). Psychological predictors of goal orientations in youth soccer. *Research Quarterly for Exercise & Sport, 65,* 355-362.

Fox, K. R., & Corbin, C. B. (1989). The physical self-perception profile: development and preliminary validation. *Journal of Sport & Exercise Psychology, 11*(4), 408-430.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology, 93,* 104-113

Graham, J. W., & Hoffer, S. M. (2000). Multiple imputation in multivariate research. In T. D. Little, K. U. Schnable and J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 201-218). Mahwah, NJ: Erlbaum.

Hagger, M. S., Chatzisarantis, N. L. D., & Biddle, S. J. H. (2002). A meta-analytic review of the theories of reasoned action and planned behavior in physical activity: predictive validity and the contribution of additional variables. *Journal of Sport & Exercise Psychology, 24,* 3-32.

Harter, S. (1992). The relationship between perceived competence, affect, and motivational orientation with the classroom: Processes and patterns of change. In A. K. Boggiano & T. S. Pittman (Eds.), *Achievement and motivation; A social-developmental perspective* (pp. 77- 113). Cambridge University Press.

Heyman, G. D., & Dweck, C. S. (1992). Achievement goals and intrinsic motivation: Their relation and their role in adaptive motivation. *Motivation and Emotion, 16,* 231-247.

Joreskog, K. G. (1979). Statistical estimation of structural models in longitudinal investigations. In J. R. Nesselroade & B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 303- 351). New York: Academic Press.

Joreskog, K. G., & Sorbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language.* Chicago: Scientific Software International.

Kaplan, D. (2000). *Structural Equation Modeling: Foundations and Extensions.* Newbury Park, CA: Sage.

Lepper, M. R. (1988). Motivation considerations in the study of instruction. *Cognition and Instruction, 5,* 289-309.

Lepper, M. R. & Sethi, A. Dialdin, D., Drake, M. (1996). Intrinsic and extrinsic motivation: A developmental perspective. In S. S. Luthar, J. A. Burack, D. Cicchetti, & J. R. Weisz. (Eds.), *Developmental psychopathology: Perspectives in adjustment, risk and disorder* (pp. 23-50). Cambridge University Press.

Lintunen, A., Valkonen, A., Leskinen, E., & Biddle, S. J. H. (1999). Predicting physical activity intentions using a goal perspectives approach: a study of Finnish youth. *Scandinavian Journal of Medicine & Science in Sports, 9,* 344-352.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: Wiley.

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to Early-adulthood. *Journal of Educational Psychology, 81,* 417-430.

Marsh, H. W. (1990). A multidimensional, hierarchical self-concept: Theoretical and empirical justification. *Educational Psychology Review. 2,* 77-171.

Marsh, H. W. (1994b). Sport motivation orientations: Beware of the jingle-jangle fallacies. *Journal of Sport and Exercise Psychology, 16,* 365-380.

Marsh, H. W. (1997). The measurement of physical self-concept: A construct validation approach. In K. Fox (Ed.), *The physical self-concept: From motivation to well-being* (pp. 27-58). Champaign, IL: Human Kinetics.

Marsh, H. W. (2002). A multidimensional physical self-concept: A construct validity approach to theory, measurement, and research. *Psychology: The Journal of the Hellenic Psychological Society, 9,* 459-493.

Marsh, H. W. (In press). Application of confirmatory factor analysis and structural equation modeling in sport/exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of on sport psychology* (3rd ed.). New York: Wiley.

Marsh, H. W., Balla, J. R., & Hau, K-T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical processes. In G. A. Marcoulides & R. E. Schumacker (Ed.), *Advanced structural equation modeling techniques* (pp. 177-198). Hillsdale, NJ: Erlbaum.

Marsh, H. W., Craven, R. G., & Debus, R. (1998). Structure, stability, and development of young children's self-concepts: A multicohort-multioccasion study. *Child Development, 69,* 1030-1053.

Marsh, H. W., Craven, R. G., Hinkley, J. W., & Debus, R. L. (2003). Evaluation of the Big-Two-Factor Theory of academic motivation orientations: An evaluation of jingle-jangle fallacies. *Multivariate Behavioral Research, 38,* 189-224.

Marsh, H. W., Ellis, L., Parada, L., Richards, G. & Heubeck, B. G. (2005). A short version of the Self Description Questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment, 17,* 81-102.

Marsh, H. W., & Hau, K-T. (1996). Assessing goodness of fit: Is parsimony always desirable? *Journal of Experimental Education, 64,* 364-390.

Marsh, H.W., Hey, J., Roche, L.A., & Perry, C. (1997). Structure of physical self-concept: Elite athletes and physical education students. *Journal of Educational Psychology, 89,* 369-380.

Marsh, H.W., Papaioannou, A., Martin, A., Theodorakis, Y. (in press). Motivational constructs in Greek physical education classes: Factor structure, gender and age effects in a nationally representative longitudinal sample. *International Journal of Sport and Exercise Psychology*

Marsh, H.W., Papaioannou, A., Theodorakis, Y. (in press). Causal ordering of physical self-concept and exercise behavior: Reciprocal effects model and the influence of physical education teachers. *Health Psychology.*

Marsh, H. W.; Wen, Z.; Hau, K. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods, 9,* 275-300.

Martin, A. J. (2001). The Student Motivation Scale: A tool for measuring and enhancing motivation. *Australian Journal of Guidance and Counselling, 11,* 1-20.

Martin, A. J. (2003). The Student Motivation Scale: Further testing of an instrument that measures school students' motivation. *Australian Journal of Education, 47,* 88-106.

Martin, A. J. Marsh, H. W., & Debus, R. L. (2001a). A quadripolar need achievement representation of self-handicapping and defensive pessimism. *American Educational Research Journal, 38,* 583-610.

Martin, A. J., Marsh, H. W., & Debus, R. L. (2001b). Self-handicapping and defensive pessimism: Exploring a model of predictors and outcomes from a self-protection perspective. *Journal of Educational Psychology, 93,* 87-102.

Martin, A. J. Marsh, H. W., & Debus, R. L. (2003). Self-handicapping and defensive pessimism: A model of self-protection from a longitudinal perspective. *Contemporary Educational Psychology, 28,* 1-36.

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

McGinnis, J. M. (1992). The public health burden of a sedentary lifestyle. *Medicine & Science in Sports & Exercise, 24,* S196-S200.

McKenzie, T. L. (2001). Promoting physical activity in youth: focus on middle school environments. *Quest, 53,* 326-334.

Newton, M., & Duda, J. L. (1993). Elite adolescent athletes' achievement goals and beliefs concerning success in tennis. *Journal of Sport & Exercise Psychology, 15,* 437-448.

Nicholls, J. G. (1989). *The competitive ethos and democratic education.* Cambridge: Harvard University Press.

Ostrow, A. C. (1990). *Directory of psychological tests in the sport and exercise sciences.* Morgantown, WV: Fitness Information Technology.

Owens, L., & Straton, R. (1980). The development of a cooperative, competitive, and individualized learning preference scale for students. *British Journal of Educational Psychology, 50,* 147-163.

Papaioannou, A. (1994). Development of a Questionnaire to Measure Achievement Orientations in Physical Education. *Research Quarterly for Exercise and Sport, 65,* 11-20.

Papaioannou, A. (1997a). Perceptions of motivational climate, perceived competence, and motivation of students of varying age and sport experience. *Perceptual and Motor-Skills, 85,* 419-430.

Papaioannou, A. (1997b): Prediction of sport violence and attitudes towards sport violence. European *Yearbook of Sport Psychology, 1,* 113-129.

Papaioannou, A. (1998). Students' Perceptions of the Physical Education Class Environment for Boys and Girls and the Perceived Motivational Climate. *Research Quarterly for Exercise and*

Sport, 69, 267-275.
Papaioannou A. & Macdonald A. I. (1993). Goal perspectives and purposes of physical education as perceived by Greek adolescents. *Physical Education Review, 16* (1), 41-48.
Papaioannou, A., & Goudas, M. (1999): Motivational climate in physical education. In Y. Vanden Auweele, F. Bakker, S. Biddle, M. Durand, & R. Seiler (Eds.), *Psychology for physical educators* (pp. 51-68). Champaign, IL: Human Kinetics.
Papaioannou, A., Marsh, H.W., Theodorakis, Y. (2004). A multilevel approach to motivational climate in physical education and sport settings: An individual or a group level construct. *Journal of Sport and Exercise Psychology, 26,* 90-118.
Papaioannou, A., & Theodorakis, Y. (1996). A test of three models for the prediction of intention for participation in physical education lessons. *International Journal of Sport Psychology, 27,* 383-399.
Qin, Z., Johnson, D. W., & Johnson, R. T. (1995). Cooperative versus competitive efforts and problem solving. *Review of Educational Research, 65,* 129-144.
Roberts, G. C. (1993). Motivation in sport; Understanding and enhancing the motivation and achievement of children. In R. N. Singer, M. Murphey, & L. K. Tennant (Eds.) *Handbook of research on sport psychology* (pp. 405-420). New York: McMillan.
Rowland, T. W. (1999). Adolescence: a risk factor for physical activity. *The Presidents Council on Physical Fitness and Sports Research Digest* (Series 3. No. 6). Washington, DC: President's Council on Physical Fitness and Sports.
Ryan, R. & Deci, E. (1989). Bringing the research traditions of task/ego involvement and intrinsic/extrinsic motivation: comment on Butler—(1987). *Journal of Educational Psychology, 81,* 265-268.
Sallis, J. F., Prochaska, J. J., & Taylor, W. C. (2000). A review of the correlates of physical activity of children and adolescents. *Medicine & Science in Sports & Exercise, 32,* 963-975.
Schutz, R. W. & Gessaroli, M. E. (1993). Use, misuse, and disuse of psychometrics in sport psychology research. In R. N. Singer, M. Murphey, & L. K. Tennant (Eds.) *Handbook of research on sport psychology* (pp. 901-920). New York: McMillan.
SPSS. (1999). *Statistical Package for the Social Sciences* (version 11). IL: Chicago.
Stipek, D. J., & Mac Iver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Development, 60,* 521-538.
Theodorakis, Y. (1994). Planned behaviour, attitude strength, role identity, and the prediction of exercise behaviour. *Sport Psychologist, 8*(2), 149-165.
Thorkildsen, T. A., & Nicholls, J. G. (1998). Fifth grader's achievement orientation and beliefs: Individual and classroom differences. *Journal of Educational Psychology, 90,* 179-201.
Treasure, D. C. (2001). Enhancing young people's motivation in youth sport: An achievement goal approach. In G. C. Roberts (Ed.), *Advances in motivation in sport and exercise* (pp. 79-100). Champaign, IL, Human Kinetics.
Treasure, D., & Roberts, G. (1995). Applications of achievement goal theory to physical education: Implications for enhancing motivation. *Quest, 47,* 475-489.
Williams, L. (1998). Contextual influences and goal perspectives among female youth sport participants. *Research Quarterly for Exercise & Sport, 69,* 47-57.

Appendix
Items Used in the Present Investigation

Task-involving Climate (7 items): T1 α = .77; T2 α = .81)

In this Physical Education Class ...

The P.E. teacher pays special attention to whether my skills are improving

The P.E. teacher looks completely satisfied when students are improving after trying hard

The P.E. teacher is completely satisfied when every student's skills are improving

The P.E. teacher is most satisfied when every student learns something new

The P.E. teacher makes sure that I understand how to perform each new skill before the class moves on to learning other skills

The P.E. teacher insists that students' mistakes are part of learning

The way the lesson is taught helps me learn how to exercise by myself

Ego-involving Climate (6 items) : T1 α = .83; T2 α = .85)

In this Physical Education Class ...

The P.E. teacher attends the best records only

The students are encouraged to play better than their schoolmates

The P.E. teacher praises the students only when they are better than their schoolmates

Only the students with the best records are rewarded

The P.E. teacher praises the students when they outperform their schoolmates

The P.E. teacher boosts the competition among the students

Task Goal Orientation: T1 α = .83; T2 α = .85)

I feel most successful in physical education lessons when ...

Something I learn makes me want to go and practice more

A skill I learn really feels right

I do my very best

I work really hard

I learn a new skill and it makes me want to practice more

I learn something that is fun to do

I learn a new skill by trying hard

Ego Goal Orientation: T1 α = .75; T2 α = .81)

I feel most successful in physical education lessons when ...

I can do better than my friends

The others can't do as well as me

I'm the best

I'm the only one who can do the play or skill

Others mess up and I don't

I score the most points

Intrinsic Motivation: Enjoyment : T1 α = .84; T2 α = .89)

In this physical education class ...

I enjoy the PE lesson very much

The PE lesson is fun

The PE lesson is very interesting

Intrinsic Motivation: Effort: T1 α = .83; T2 α = .86)

In this physical education class ...

I put high effort in the PE lesson

It is important to me to do well in the PE lesson

I try hard while I am practicing in the PE lesson

Exercise Attitudes: T1 α = .58; T2 α = .75)

Doing regular exercise in the next 12 months is ...

very good = 7, very bad = 1,

very healthy = 7, very unhealthy = 1,

very useful = 7, very useless = 1

Perceived Behavioral Control: T1 α = .83; T2 α = .90)

For me, doing regular exercise in the next 12 months is (very easy = 7, very difficult = 1)

I can exercise regularly in the next 12 months (very possible = 7, very impossible = 1)

I am absolutely certain that I will exercise regularly in the next 12 months (absolutely right = 7, absolutely wrong = 1)

continued

Appendix
Items Used in the Present Investigation *(Continued)*

Exercise Intention: T1 α = .81; T2 α = .87)

I intend to exercise regularly in the next 12 months (very possible = 7, very impossible =1)

I am determined to exercise regularly in the next 12 months (absolutely yes = 7, absolutely no =1)

Exercise behavior:

How many times did you exercise in the last month? None, 1-5, 5-10, 10-15, 15-20, over 20

Physical Self-concept: T1 α = .80; T2 α = .82)

Some people feel that they are good when it comes to playing sports

Some people feel that they are among the best when it comes to athletic ability

Some people are quite confident when it comes to taking part in sports activities

Some people feel that they are always one of the best when it comes to joining in sports activities

Given the chance, some people are always one of the first to join in sports activities

Note. T1 α and T2 α refer to coefficient alpha estimates of reliability at Times 1 and 2.

Acknowledgments

We would like to acknowledge financial assistance for this research provided by the Greek Ministry of Education (Center of Educational Research) and to offer thanks to the many colleagues from the University of Thrace and the University of Thessaly who collaborated in the planning and implementation of the project, and to Joan Duda and Ken Fox who were external advisers to the project. In addition, we would like to acknowledge the Australian Research Council's support of Professor Marsh's involvement in this research. Importantly, we would also like to thank the many physical education teachers and students who graciously volunteered to participate in the research.