Cross-Cultural Validation of the Task and Ego Orientation in Sport Questionnaire

Fuzhong Li
Oregon Research Institute

Peter Harmer
Willamette University

Likang Chi
National College of Physical Education and Sport

Naruepon Vongjaturapat
Burapha University

It is becoming increasingly important to determine whether structural models of measures of sport and activity behavior developed in North America are invariant across different populations. This study assessed (a) the cross-cultural validity of the Task and Ego Orientation in Sport Questionnaire (TEOSQ) using male college students across the United States ($n = 309$), Thailand ($n = 312$), and Taiwan ($n = 307$); and (b) the factorial equivalence and structured latent mean differences of the TEOSQ in these samples. Using a confirmatory factor analytic procedure, the initial test of the hypothesized two-factor structure representing task and ego orientation yielded a good fit for each sample. The factor structure was further shown to be metric invariant across the three countries. Furthermore, tests of latent means showed significant differences between groups. The United States sample exhibited the highest levels of task and ego orientation, followed by the Taiwan and Thailand samples, respectively.

Key words: goal orientation, cross-cultural comparisons, confirmatory factor analysis, metric invariance

As psychological theories and measures pertaining to sport and physical activity accumulate in the United States, so do questions concerning their applicability to other countries and cultures, especially as researchers attempt to use these theories and measures for cross-cultural comparative studies (Duda & Allison, 1990). Thus, examining whether theories and measures developed in the United States are applicable in other countries is becoming increasingly important. Fundamental to this process is establishing the invariance of the factor structure of these theories and measures.
measures across different populations. Without this invariance, findings based on these measures are not meaningful. Therefore, this study was designed to examine the cross-cultural validity of goal perspective theory (Nicholls, 1989) and its corresponding measure, the Task and Ego Orientation in Sport Questionnaire (TEOSQ; Duda & Nicholls, 1992). To do so, we examined the equivalence issues of sport goal orientation construct and substantive predictions about the mean differences on the two achievement goal orientations within the domain of cross-cultural research.

Achievement Goal Orientation Theory

Nicholls's (1989) goal perspective theory proposes that two goal orientations, task orientation and ego orientation, operate in achievement contexts. Nicholls's theory states that individuals strive to demonstrate competence through task or ego goal orientations. According to the theory, task orientation implies task mastery or personal improvement as reflecting high competence and subjective success. Thus, perception of competence tends to be self-referenced if the individual is task oriented. Ego-oriented individuals, on the other hand, tend to gauge their own competence by comparison to others. In these individuals, high ability and perceived goal accomplishment are predicated on displaying superiority over those to whom they compare themselves. Achievement goal orientation theories assert that an individual’s goal perspective will affect self-evaluations of demonstrated ability, expended effort, and attributions for success and failure. In turn, these cognitions are assumed to influence achievement-related affect and subsequent behaviors such as performance, task choice, and persistence (see Duda, 1992, 1993, for a review).

Within the theoretical framework of achievement goal orientation, Duda and Nicholls (1992) developed the TEOSQ to assess individual differences in proneness for task or ego goal orientation in athletic settings. Psychometric tests on TEOSQ have shown some construct validity, as well as internal consistency reliability, in research involving a variety of samples, including youth sport, high school students, and college-age sport participants and nonparticipants in North America (e.g., Chi & Duda, 1995; Duda & Nicholls, 1992; Li, Harmer, & Acock, 1996; White & Duda, 1994).

The availability of the TEOSQ measure has led to an on-going research effort in understanding motivational implications of holding goal perspectives in sport and physical activity. Studies using the TEOSQ have examined both the cognitive correlates and predictors of dispositional goal orientations in athletic settings (Duda, 1992, 1993; Roberts, 1993). The extant research suggests that both situational (e.g., perceptions of motivational climate, situationally induced goals) and individual-difference factors (e.g., task vs. ego orientation) influence achievement goal orientation. Moreover, research indicates that the two goal orientations affect adaptive motivational behaviors (e.g., intrinsic motivation, attributions, athletic performance).

Cross-Culture Theory Testing

The development of the TEOSQ has facilitated empirical research on achievement goal orientation in the field of sport psychology. However, relatively little research has been conducted in the cross-cultural context (e.g., Duda, Fox, Biddle, &
Armstrong, 1992; Hayashi & Weiss, 1994; Kim, 1995; Li, Harmer, Acock, Vongjaturapat, & Boonverabut, in press; Williams, Kim, & Gill, 1995). In particular, little has been done to determine whether the TEOSQ measure operationalizes goal orientation construct equivalently across diverse cultural groups. Both the theory and its measure have been developed and tested exclusively in North America, and the extent to which they may be culturally biased remains largely unknown. In practice, there is an implicit assumption that theories and measures developed in North America are relevant and applicable to other countries and cultures, without being subjected to an actual cross-cultural validation process. This assumption can lead to invalid cross-cultural inferences. In this regard, there is a growing need to develop constructs and measures that are reliable, valid, and applicable across cultures. Measures with sound psychometric properties across countries will facilitate cross-cultural comparative research and should help in understanding diverse culture variations in values, orientations, and motivation of sport and physical activity participation around the world. Although researchers have called for an examination of the applicability of theories of sport psychology to other countries and cultures (Duda & Allison, 1990), few studies have actually done so.

The cross-cultural psychology literature suggests that both conceptual (i.e., the meaning of a construct) and metric (i.e., the scalar of an instrument) equivalence must be established before any cross-cultural comparisons or conclusions regarding the generalizability of a theory can be made (Hui & Triandis, 1985). In many domains of psychological research, the issue of measurement invariance of a particular factor structure is also viewed as a necessary prerequisite for testing mean differences on constructs across groups (e.g., Byrne, Shavelson, & Muthén, 1989). A central principle in multisample or cross-cultural studies is that for any psychological measurement the relationship between the trait (e.g., task construct) and its corresponding indicators (e.g., measurement items of the task) should be invariant across subpopulations. In other words, the measurement items should be equally valid across groups (i.e., the perception of item content should not vary across groups).

To date, most cross-cultural research on goal orientation in sport that involved examination of group means has not considered the factorial equivalence of TEOSQ structure across cultures (e.g., Hayashi & Weiss, 1994; Williams, Kim, & Gill, 1995). Those validation studies that have examined the measurement equivalence issue have, however, relied on the use of a single nation sample (e.g., Chi, 1994; Kim, 1995; Li et al., in press) and, as such, provided no direct cross-national empirical tests on the invariance assumptions. Therefore, there are both conceptual and methodological necessities for testing TEOSQ metric equivalence cross-culturally. Conceptually, it is imperative that the proposed goal perspective be examined cross-culturally in order to extend our understanding of motivational orientation in sport participation. Methodologically, demonstrating construct validity, as well as equivalence, cross-culturally would facilitate substantive research that focuses on cultural comparisons of achievement goal orientation.

The present investigation extends the existing cross-cultural construct validation research by examining the psychometric properties of TEOSQ for samples from the United States, Thailand, and Taiwan. Specifically, the purposes of the present study were twofold: (a) to test the validity of a two-factor TEOSQ measurement model for samples from the United States, Thailand, and Taiwan, and (b) to examine the factorial equivalence of the TEOSQ and its structured latent mean differences in levels of goal orientation across these samples.
Method

Participants

The participants in the study were male college students recruited from university-based physical education activity classes. Of the total participant sample ($N = 928$), 309 (age, $M = 20.98$ years, $SD = 1.65$) were from the United States, 312 (age, $M = 21.74$ years, $SD = 1.87$) were from Thailand, and 307 (age, $M = 20.43$ years, $SD = 1.98$) were from Taiwan. The United States sample was predominantly Caucasian, while the Thailand and Taiwan samples were culturally homogeneous. Participants were engaged in physical activity skills classes including basketball, martial arts, soccer, volleyball, and weight training. Because participants were college students involved in sport-related activities, the samples were considered equivalent for both activities and the level of instruction to which they were exposed.

Measure

The 13-item TEOSQ (Duda & Nicholls, 1992) was used in this study. Although the TEOSQ has been used most frequently with competitive sports participants, the underlying psychological constructs are considered to be shared by individuals involved in any achievement-related physical activity (Duda, 1992). An English version of the TEOSQ was administered to the United States sample. A translated Thai (Li et al., in press) and Chinese (Chi, 1994) version were used for the Thailand and Taiwan samples, respectively. To ensure meaning equivalence of TEOSQ items in these samples, the back translation procedure outlined by Brislin (1986) was employed. Specifically, the TEOSQ was first translated into Thai and Chinese by bilingual language experts in the respective countries, and then re-translated back into English by different bilingual language experts in the United States to uncover discrepancies. No discrepancies were found in the final versions of the TEOSQ used in the present study.

Procedures

Protocol and procedures for this study were approved by the Institutional Review Board (IRB) of an American university. Guidelines for human subjects protection vary in the other countries involved in this study. Where discrepancies may have become problematic, the higher standard of the initial IRB guidelines were followed. After obtaining permission from both program coordinators and from the instructors within each country, data were collected during a scheduled class period. Participants were informed that the purpose of the study was to investigate reasons for sport participation, that participation in the study was voluntary, and that anonymity was guaranteed. Return of a completed questionnaire was accepted as indicating the individual’s consent to participate.

The a Priori Model Specifications

An a priori two-factor model representing task and ego orientation in sport was specified separately for each country. The model was specified in accordance with goal perspective theory in that the two goal orientations were assumed to be
independent (Nicholls, 1989). The TEOSQ measurement model representing each sample is shown in Figure 1. In Figure 1, we followed the suggestion by MacCallum (1995) and used a simplified path model presentation to illustrate our TEOSQ measurement model specification. In the figure, ellipses represent latent variable (LV) of task and ego. Because both LVs are considered exogenous in this measurement model, the variance of LV is set at a value of 1.0 to establish a scale. The asterisk represents values to be freely estimated. The first category of such values is associated with the single-headed paths from the LVs to indicators (TEOSQ items in rectangles) and represents unknown factor loadings. Each indicator in the rectangle is assumed to contain measurement error with its own unique variance (the second group of values to be freely estimated). These error terms are hypothesized to influence corresponding indicators, with an associated weight of 1.0.

After fitting individual sample models, they were stacked into a multisample analysis that allowed examination of the hypothesis of equivalent TEOSQ factor structure, particularly the factor loadings, across the three countries. In this regard, two models were specified. The first model involved testing the hypothesis of equivalent factor form (i.e., Task and Ego). The second model hypothesized equivalent factor loadings (i.e., metric equivalence). Final model specification focused on structured latent means. This included models of (a) equivalent intercepts, and (b) equivalent structured latent means across the three samples. It should be noted that a priori model specifications allow researchers to engage in model testing within a hypothesis-testing framework.

**Data Analysis**

Models were estimated via confirmatory factor analysis (CFA) using the LISREL program (Jöreskog & Sörbom, 1993a). As the data were ordinal, asymptotic covariance matrices were generated with PRELIS (Jöreskog & Sörbom, 1993b), a preprocessor of LISREL, where polychoric correlations are assumed between ordinal variables. Estimation and tests of models were then performed with LISREL where the weighted least squares (WLS) method was employed. The large sample size required by the procedure was met in the present study.¹

Data were analyzed via CFA procedures at both single (national level) and multiple-group (cross-national) levels in three stages. The first stage involved establishing, at a national level, a baseline TEOSQ measurement model for each country. In addition, the model based on the pooled data was analyzed using a technique for “deculturing” the data (see the discussion in the Results section). In the second stage, CFA approaches to factorial invariance were used to test, at a multinational level, the equality of parameter estimates across the three countries. In the final stage of analyses, country differences in structured latent means were examined.

In testing models concerning the invariance of latent factor structure, there are four major hypotheses of interest: (a) factor form, (b) factor loadings, (c) variance/covariance, and (d) error/uniquenesses (Bollen, 1989; Byrne, 1989). Because the primary focus of this study was to establish cross-culturally the TEOSQ metric equivalence, emphasis was placed on testing the first two key hypotheses: invariance of factor form and factor loadings. Evidence for the support of these two hypotheses indicates TEOSQ metric equivalence across the samples.

Because of the problems associated with the traditional chi-square statistic test (Marsh, Balla, & McDonald, 1988), we adopted multiple indices in our model.
Figure 1 — TEOSQ measurement model representing the United States, Thailand, and Taiwan samples. t1–t7 = items for Task subscale; e1–e6 = items for Ego subscale; E1–E13 = measurement errors.
evaluation: Comparative Fit Index (CFI; Bentler, 1990) and Tucker-Lewis Index (TLI; Tucker & Lewis, 1973). In testing for model parameter invariance across samples, we incorporated a newly proposed index: Hierarchical TLI (HTLI; Marsh, 1994). Thus, although the chi-square test statistic is presented, greater emphasis is placed on the later three indices in our global model evaluation.

Results

Descriptive Statistics

Inspection of the means for the two TEOSQ subscales showed higher levels of responses on both Task ($M = 4.193$, $SD = .463$) and Ego ($M = 2.783$, $SD = .877$) for the United States sample. This was followed by the Taiwan sample (Task, $M = 4.060$, $SD = .456$; Ego, $M = 2.546$, $SD = .673$). The Thailand sample showed the lowest responses in both subscales (Task, $M = 3.398$, $SD = .615$; Ego, $M = 2.221$, $SD = .725$). The differences in TEOSQ subscales across the samples may be viewed as a preliminary indicator that the three samples differ on the levels of sport goal orientation.

Tests of multivariate skewness and kurtosis from PRELIS indicated significance: skewness $= 35.711$ ($p < .001$) for the United States sample, 28.719 ($p < .001$) for the Thailand sample, and 29.061 ($p < .001$) for the Taiwan sample; kurtosis $= 17.656$ ($p < .001$) for the United States sample, 14.795 ($p < .001$) for the Thailand sample, and 14.079 ($p < .001$) for the Taiwan sample. These results suggested a nonnormal multivariate distribution and thus justified the use of asymptotically distribution-free analysis with LISREL WLS method (Jöreskog & Sörbom, 1993a).

National Level: Single-Sample Analysis

The national-level analysis assessed whether the psychometric properties of TEOSQ (i.e., dimensionality and internal consistency) exhibited a similar pattern across nations. CFA suggested a fairly good fit of the a priori TEOSQ measurement model for each country. Both TLI and CFI values exceeded .90. These results were encouraging because it is important to establish a well-fitting and theoretically viable baseline measurement model before any constraints can be imposed in a cross-multiple sample analysis. Fit statistics for each of the three models are shown in the top portion of Table 1. Notice that all of the models could be rejected on a statistical basis (see the overall chi-square statistics in the second column). This is not uncommon when large sample sizes are analyzed.

Parameter estimates from LISREL WLS estimation for each measurement model are presented in Table 2. Inspection of the unstandardized factor loadings, shown in the second through fourth columns of Table 2, indicated a moderate relationship between observed indicators (i.e., TEOSQ items) and hypothesized latent factors (i.e., Task and Ego). All factor loadings were significant ($p < .001$). Also, inspection of the squared multiple correlations for the observed indicators revealed that the latent factors explained a moderate amount of the item variance across the three samples: $M = .570$ for the United States sample, $M = .482$ for the Thailand sample, and $M = .446$ for the Taiwan sample.

Factor loadings appear quite similar across the three samples ($M = 1.417$ for the United States sample, $M = 1.316$ for the Thailand sample, and $M = 1.297$ for the Taiwan sample). To facilitate parameter estimates comparison across groups, a
## Table 1  Model Fitting Indices Across the Three Countries

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>$df$</th>
<th>$\chi^2_{\text{diff}}$</th>
<th>$df_{\text{diff}}$</th>
<th>TLI</th>
<th>CFI</th>
<th>HTLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test the baseline model for each country</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>386.52*</td>
<td>65</td>
<td>—</td>
<td>—</td>
<td>.977</td>
<td>.981</td>
<td>—</td>
</tr>
<tr>
<td>Thailand</td>
<td>200.92*</td>
<td>65</td>
<td>—</td>
<td>—</td>
<td>.967</td>
<td>.972</td>
<td>—</td>
</tr>
<tr>
<td>Taiwan</td>
<td>239.53*</td>
<td>65</td>
<td>—</td>
<td>—</td>
<td>.936</td>
<td>947</td>
<td>—</td>
</tr>
<tr>
<td>Total sample</td>
<td>490.70</td>
<td>65</td>
<td>—</td>
<td>—</td>
<td>.941</td>
<td>.951</td>
<td>—</td>
</tr>
<tr>
<td>Tests of invariance of model across the three countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pattern (Model 1)</td>
<td>826.97</td>
<td>195</td>
<td>—</td>
<td>—</td>
<td>.969</td>
<td>.975</td>
<td>1.00</td>
</tr>
<tr>
<td>Loading (Model 2)</td>
<td>951.80</td>
<td>217</td>
<td>124.83*</td>
<td>22</td>
<td>.968</td>
<td>.973</td>
<td>.989</td>
</tr>
<tr>
<td>Tests of structured mean differences across the three countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Model 3)</td>
<td>1,085.07</td>
<td>239</td>
<td>133.27*</td>
<td>22</td>
<td>.962</td>
<td>.967</td>
<td>.977</td>
</tr>
<tr>
<td>Latent (Model 4)</td>
<td>4,222.28</td>
<td>243</td>
<td>2,987.21*</td>
<td>5</td>
<td>.849</td>
<td>.843</td>
<td>0</td>
</tr>
</tbody>
</table>

*Note.* TLI = Tucker and Lewis Index; CFI = Comparative fit Index; HTLI = Hierarchical TLI. *p < .001.

## Table 2  Estimated Parameter Matrices for the Single Sample, Across Samples, TEOSQ Measurement Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>United States estimate</th>
<th>Thailand estimate</th>
<th>Taiwan estimate</th>
<th>Across three countries*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>1.000*</td>
<td>1.000</td>
<td>1.000</td>
<td>0.771 (.028)*</td>
</tr>
<tr>
<td>Task 2</td>
<td>1.122 (.059)</td>
<td>1.067 (.043)</td>
<td>1.110 (.058)</td>
<td>0.654 (.030)</td>
</tr>
<tr>
<td>Task 3</td>
<td>1.816 (.035)</td>
<td>1.579 (.054)</td>
<td>1.612 (.068)</td>
<td>0.645 (.012)</td>
</tr>
<tr>
<td>Task 4</td>
<td>1.795 (.054)</td>
<td>1.499 (.067)</td>
<td>1.500 (.073)</td>
<td>0.590 (.031)</td>
</tr>
<tr>
<td>Task 5</td>
<td>1.335 (.045)</td>
<td>1.404 (.045)</td>
<td>1.304 (.058)</td>
<td>0.661 (.023)</td>
</tr>
<tr>
<td>Task 6</td>
<td>1.140 (.055)</td>
<td>1.440 (.057)</td>
<td>1.039 (.063)</td>
<td>0.676 (.028)</td>
</tr>
<tr>
<td>Task 7</td>
<td>1.855 (.063)</td>
<td>1.569 (.055)</td>
<td>1.533 (.072)</td>
<td>0.607 (.030)</td>
</tr>
<tr>
<td>Ego 1</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.728 (.034)</td>
</tr>
<tr>
<td>Ego 2</td>
<td>1.278 (.060)</td>
<td>1.383 (.052)</td>
<td>1.261 (.057)</td>
<td>0.661 (.026)</td>
</tr>
<tr>
<td>Ego 3</td>
<td>1.235 (.068)</td>
<td>1.347 (.063)</td>
<td>1.335 (.063)</td>
<td>0.767 (.029)</td>
</tr>
<tr>
<td>Ego 4</td>
<td>1.486 (.062)</td>
<td>1.239 (.054)</td>
<td>1.386 (.064)</td>
<td>0.624 (.029)</td>
</tr>
<tr>
<td>Ego 5</td>
<td>1.216 (.066)</td>
<td>0.989 (.051)</td>
<td>1.063 (.074)</td>
<td>0.610 (.028)</td>
</tr>
<tr>
<td>Ego 6</td>
<td>1.303 (.074)</td>
<td>0.963 (.064)</td>
<td>1.120 (.076)</td>
<td>0.754 (.031)</td>
</tr>
</tbody>
</table>

*Note.* Values in the parenthesis are standard error estimate.

*This is a solution based on common metric standarized estimates (Jöreskog & Sörbom, 1989). *Parameters fixed at tabled values to identify each model. *This is the rescaled standard errors (SE) computed using the equation: standardized SE = standarized estimate/t value (cf. Marsh, 1993, p. 265).
solution based on common metric standardized estimates (Jöreskog & Sörbom, 1989, p. 238) is presented (shown in the fifth column of Table 2). These loadings (ranging from .590 to .771, $M = .673$) suggest a moderate items variance explained.

Because goal perspective theory (Nicholls, 1989) assumes that the two goal orientations (task and ego) are independent (as shown in Figure 1), an empirical examination of this relationship is of theoretical interest. The relationship was estimated by adding an additional path linking the two constructs. The result showed that, consistent with previous studies (Li & Harmer, 1993; Li et al., 1996), there was a nonsignificant relationship between task and ego orientation ($\phi_{21} = -.028, t = -1.288, p > .384$, where $\phi_{21}$ stands for the correlation between the two latent constructs) for the United States sample. A nonsignificant relationship was also found for the Taiwan sample ($\phi_{21} = .029, t = 1.836, p > .067$). However, the relationship was found to be statistically significant and positively correlated for the Thailand sample ($\phi_{21} = .393, t = 4.790, p < .001$), indicating that a high level of task orientation was associated with a high level of ego orientation. This result was compatible with the finding based on a Korean sample ($r = .28$, where $r$ stands for zero-order correlation; Kim, 1995) but incompatible with those obtained from a Thailand intercollegiate athlete male sample ($\phi = -.84$; Li et al., in press).

The internal consistency estimate (alpha reliability, Cronbach, 1951) was computed separately for the two subscales of the TEOSQ across the three samples. Estimates showed that across the two subscales, the United States sample ($\alpha = .867$ for Task, .897 for Ego) tended to provide more reliable responses than the Thailand sample ($\alpha = .862$ for Task, .838 for Ego) or the Taiwan sample ($\alpha = .800$ for Task, .804 for Ego). The differences, however, were small and consistent across the three countries. In sum, the TEOSQ dimensionality tests and internal consistency estimates indicated that the goal orientation measure is valid and reliable within each country. Moreover, there appeared to be a similar pattern in the magnitude of the parameter estimates (i.e., factor loadings) across the three samples.

**Pooled Data Analysis**

To further evaluate the cross-cultural applicability of the model, the data were pooled and analyzed with a method proposed by Bond (1988). With this method, the data are first “decultured” across samples by removing culturally idiosyncratic patterns from the data. This involves standardizing each observed variable within each country’s sample separately. In this manner, the true correlation between any two variables is unaffected by culture-specific factors because the average score for the variables in each sample is zero. The decultured data are then pooled across countries and analyzed in aggregate.

The two-factor model again provided an adequate fit to the data (see Total Sample in Table 1). Consistent with the national level analysis, the factor loadings (standardized) were moderately high, ranging from 0.612 to 0.859 for Task and 0.654 to 0.870 for Ego, respectively. All loadings were statistically significant ($p < .001$).

**Cross-National Level: Multisample Analysis**

The multigroup analysis offers a stringent test of the two-factor goal orientation model in terms of its cross-cultural comparability. In the present investigation, we focused on assessment of the metric equivalence that involved testings for equivalent factor form and factor loadings of the TEOSQ factor structure across the United States, Thailand, and Taiwan samples.
The first test, Model 1, of the comparability of TEOSQ factor structure across samples concerns evaluating the extent to which the factor forms are similar. This model essentially imposed no equality constraints across samples (i.e., a totally free model). The result of this analysis showed that the model fit the observed data well (see fit indices of TLI and CFI, in the middle portion of Table 1). Notice that the HTLI is 1.000 since this model represents a baseline for comparing all subsequent models that involve invariant constraints on various parameter estimates and defines one endpoint for the hierarchical relative fit indices (Marsh, 1994). Thus, this test indicated the hypothesis of equivalent factor form was tenable and the two-factor theoretical structure was considered to represent an adequate fit to the data.

The next model, Model 2, which tested the metric invariance hypothesis, imposed equality constraints on the free factor loadings across countries to assess the degree to which the scaling of the TEOSQ measurement model in the three samples was invariant. The fit of the model was again satisfactory as indexed by TLI, CFI, and HTLI (middle portion of Table 1); fit indices were nearly identical to those of Model 1. This indicated that the imposition of identical factor loadings does not make the fit substantively worse, at least from a practical standpoint. The applicability of such constraints indicates that the goal orientation indicators are performing their measurement task equally well across the three countries.

The tests conducted to this point provided evidence of the comparability of the two-factor model among the three countries. It should be emphasized that without evidence of measurement invariance for these samples, comparisons of differences in their levels of task and ego orientation would lack clear meaning. Since metric invariance had been demonstrated, it was sensible to analyze sample differences in levels of the two types of goal orientation.

**Structured Latent Means**

As with previous invariance analyses, examination of structured means models involved testing simultaneously across the three samples. The analysis required imposing constraints on both the intercepts of the observed variables and means of the latent constructs in addition to restrictions on the slopes (i.e., factor loadings) of these observed variables.

The first test of structured latent means, Model 3, involved imposing restrictions on the intercepts (i.e., establishing equivalent intercepts across the three samples). This restriction did not worsen the fit although there was a slight decline in TLI, CFI, and HTLI indices (see bottom of Table 1). The change from the invariance of the slopes model (i.e., Model 2) to the invariance of the intercepts model is considered small (ΔTLI = .007, ΔCFI = .003, and ΔHTLI = .012, where Δ is the difference in the fit index between the Model 2 and Model 3 test) and suggests that the equality of intercepts can be accepted from a practical standpoint. Results thus provided support for the invariant intercept hypothesis suggesting that the slopes and intercepts for the TEOSQ items do not vary across countries. In substantive terms, model invariance at this level indicated that across the three samples, TEOSQ items measured the two latent constructs (task and ego) equally well and participants from these samples showed systematic responses on these items.

The next test, Model 4, restricted the structured latent means to be invariant across the three samples. The model revealed an apparent distortion of fit: All practical fit indices (the last row of Table 1) dropped substantially, and all values were
below the conventional acceptable level. This implied that significant mean differences existed for the task and ego constructs across the countries. Follow-up tests examining specific group mean differences were then conducted. For identification purposes (see Bollen, 1989; Byrne, 1989), one of the samples was used as the reference group with its latent mean parameters (i.e., Task and Ego factors) fixed at zero. Mean parameters for the remaining two countries were free to differ from zero. For example, when comparing the United States sample with the Thailand and Taiwan samples, we fixed the latent means of task and ego orientation at zero for the United States sample. If one or both of these estimates for the Thailand and Taiwan samples were positive, it would indicate that the Thailand sample, Taiwan sample, or both had higher task and ego orientations than the United States sample. Alternatively, if one or both estimates for the Thailand and Taiwan samples were negative, it would indicate lower levels of goal orientation. Comparison of the group differences on latent means is based on the difference from zero. Statistical significance is determined from the \( t \) values.

Table 3 presents latent mean parameter estimates. When the United States sample was used as the reference group, results showed significant mean group differences across the three samples with respect to task and ego factors. Specifically, both the Thailand and Taiwan samples exhibited lower mean values of task and ego orientation than the United States sample. For the Thai sample, task \( t = -26.042, p < .001 \); ego \( t = -14.588, p < .001 \). For the Taiwan sample, task \( t = -12.667, p < .001 \); ego \( t = -11.706, p < .001 \). The negative \( t \) value indicates that the parameter estimates in the nonreference group samples have a significantly lower value than in the reference group.

Because the above test does not indicate the difference in latent means between the Thailand and Taiwan samples, an additional test in which the Thailand sample was used as a reference group was conducted. Results showed that the Taiwan sample had a significantly higher mean both in task (parameter estimate = .422, \( t = 20.543, p < .001 \)) and ego orientation (parameter estimate = .049, \( t = 2.769, p < .001 \)) than the Thai sample.

In addition to the evaluation of the size of mean differences in relation to standard error, further comparisons were made by standardizing the latent means. Estimates are presented in Table 3. The standardization was accomplished by dividing each latent mean estimate by its standard deviation (i.e., the square root of the latent variance) so that comparisons of latent mean differences across the samples were made in a standardized form (Marsh & Grayson, 1990). Differences were

### Table 3 Structured Mean Differences

<table>
<thead>
<tr>
<th>Construct</th>
<th>United States</th>
<th>Thailand Unstand. ( M )</th>
<th>Thailand Stand. ( M )</th>
<th>Taiwan Unstand. ( M )</th>
<th>Taiwan Stand. ( M )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task</td>
<td>0.0*</td>
<td>-0.625* (0.024)</td>
<td>-0.873</td>
<td>-0.190* (0.015)</td>
<td>-0.422</td>
</tr>
<tr>
<td>Ego</td>
<td>0.0</td>
<td>-0.248* (0.017)</td>
<td>-0.541</td>
<td>-0.199* (0.017)</td>
<td>-0.369</td>
</tr>
</tbody>
</table>

*Note. Values in the parenthesis are standard error of estimates.

*The value was set to zero as a reference group.

*\( p < .001 \).
moderate to large in task and ego orientation between the United States sample and the Thailand sample and relatively small between the United States sample and the Taiwan sample. In addition, the difference in task orientation was large between the Thailand and Taiwan sample (parameter estimate = −.830), but small in ego orientation (parameter estimate = −.087). Collectively, these results suggest that latent mean structures are not equivalent across these countries, implying significant variations in perceptions about the way participants define success in different cultures.

Discussion

Many motivational theorists in sport and physical activity emphasize the importance of considering factors such as culture, race, and ethnicity (e.g., Duda & Allison, 1990). As a consequence, we are beginning to see empirical studies focusing on this issue (e.g., Duda et al., 1992; Hayashi & Weiss, 1994; Kang & Gill, 1990; Kim, 1995; Li et al., in press). To further this process, the primary objective of this study was to demonstrate the cross-cultural validity of goal perspective theory and its corresponding measure. With this focus, four procedures (a national-level analysis, a pooled-data analysis, a multisample analysis, and a structured mean analysis) were employed to examine the sport goal orientation measure (TEOSQ) and the levels of task and ego orientations across three countries.

Results obtained from national-level and pooled-data analyses were consonant with the factor structure of the TEOSQ presented by Duda and Nicholls (1992) as a two-dimensional sport goal orientation construct. Thus, support for the TEOSQ’s factor structure and reliability can be claimed for the samples from each of the countries used in this study. This suggests that the factor structure of the TEOSQ represents a valid and reliable theoretical construct for studying athletic goal orientation among college-aged male students in the United States, Thailand, and Taiwan.

An interesting result of these analyses is the relationship observed between task and ego orientation. In contrast to goal perspective theory, a positive relationship between the two goal orientations was found in the Thailand sample (i.e., a high task orientation was related to a high level of ego orientation). This finding appears consistent with Kim (1995) who conducted a TEOSQ validation study with a Korean sample (although the relationship was estimated based on zero-order correlation) but contrary in its direction to data reported by Li et al. (in press) who found a high negative relationship between task and ego in their study of Thai intercollegiate athletes. The correlation observed in the United States sample is small and of a magnitude consistent with previous results (Li & Harmer, 1993; Li et al., 1996). Taken together, these results suggest that the relationship between the two goal orientations may vary as a result of the interaction of levels of sport participation (e.g., instructional vs. competitive) and culture. As predicted by goal perspective theory, the relationship appears to be minimal or orthogonal for the United States and Taiwan samples but moderately related for the Thailand sample.

A substantive issue addressed in the present study was the evaluation of the TEOSQ factor structure simultaneously across diverse cultures. Tests for metric equivalence of the TEOSQ measurement model demonstrated the psychometric invariance of TEOSQ dimensionality and item measurement in a multicultural comparison. These results capture the theoretically attractive distinction between task and ego orientation and provide rigorous evidence of goal orientation construct validity using the TEOSQ in a cross-cultural setting. It can be argued that
the 13-item TEOSQ measurement model is equally valid for the present samples of college male students despite differences in culture. Moreover, future research of cross-cultural comparisons in the levels of achievement goal orientation among these samples is fully justified.

Implicitly assuming intergroup measurement invariance of the TEOSQ, previous cross-cultural studies have used observed scores or factor scores to compute the mean differences. However, without demonstrating that the TEOSQ operationalizes the goal orientation construct similarly across diverse cultural groups, researchers cannot place confidence in the results of the substantive tests conducted. This study adopted a more stringent procedure (i.e., structural equation modeling) to verify the TEOSQ factor structure in terms of its metric equivalence prior to the test of group mean differences. As a result, findings for differences in mean levels can be interpreted validly and meaningfully.

The results from the present study indicated that although the TEOSQ revealed invariant factorial structure, the mean values on the latent constructs of task and ego orientation showed differences across the countries. The United States sample showed the highest means on both task and ego orientation, followed in order by the Taiwan sample and the Thailand sample. These results indicate that although the underlying pattern of the theoretical construct is similar, the mean scores on the constructs can differ.

Although results pertaining to the structure of the goal orientation construct were consistent across countries, they are tempered by certain limitations. First, although the samples were matched in terms of age, education, and gender breakdown, they cannot be considered totally representative of the populations of the three countries. Additionally, this study was limited to male college students. The question of significant culturally based gender differences has not been examined. Finally, studies using samples with greater within-group heterogeneity are needed to confidently generalize the findings beyond the samples used here.

Given evidence of the validity of the TEOSQ in these cultures, several areas appear fruitful for further research. First, future studies should include examination of the nomological validity of the TEOSQ in these countries, particularly in Thailand and Taiwan. For example, future research should link the TEOSQ subscales to theoretically relevant constructs such as sport participation motivation, perceptions of athletic competence, and behavioral indicators such as effort and persistence (Duda, 1992, 1993). Second, given that the TEOSQ measures similar underlying construct in these groups, future research could examine differences and similarities in correlates and predictors of goal orientation in these groups as it may illuminate the context and meaning of dimensions of goal orientation across the diverse cultural groups that participate in sport (Duda, 1993). Finally, results from the present investigation suggest that translating the TEOSQ and applying it to more countries that vary in social and cultural background should enhance our understanding of the psychology of sport participation, in particular, the dispositional goal perspective in sport.

In conclusion, the usefulness of the TEOSQ for international comparative studies holds promise. Given the growing trend of cross-cultural comparative studies, testing theories cross-culturally is surely warranted. This study represents a first step in testing achievement goal orientation theory across diverse cultural groups. Empirical verification of this kind is critical prior to the application of any theory in a cross-cultural context.
Cross-Cultural Validation of TEOSQ

References


Notes

1Jöreskog and Sörbom (1993b) define a large sample as equal to or larger than $k(k - 1)/2$ cases, where $k$ equals the number of observed variables. Applying this formula [13(13-1)/2], 79 cases were needed from each country. However, literature on structural equation modeling applications indicates that asymptotically distribution free estimators such as WLS require much larger sample sizes in order to produce stable estimates (see West, Finch, & Curran, 1995).

2The HTLI addresses the concerns raised by Sobel and Bohrnstedt (1985) regarding the appropriateness of the traditional null model as a baseline for comparison in many model-testing situations. The formula for computing HTLI as provided by Marsh (1994, p. 34) is the following: $HTLI = [(\chi^2_{IT}/df_{IT}) - (\chi^2_S/df_S)]/(\chi^2_{IT}/df_{IT}) - (\chi^2_{TF}/df_{TF})]$, where TI = totally invariant model, TF = totally free model, and M = model of intermediate complexity being considered. Values approaching or greater than 1.0 indicate that the ratio of the difference between the TI and the TF is closer to the invariant model (i.e., the model of interest in the invariance test hierarchy).

3It is worth noting that the model would be rejected based on the chi-square difference test, $\chi^2 (22, N = 928) = 124.83, p < .001$. The test, however, like the conventional chi-square test, is complicated by its sensitivity to large sample sizes as in this study. Following the suggestion by Marsh (1994) for the use of subjective indices for evaluating the support for the invariance constraints, we decided to rely on the three practical fit indices for our model evaluation.

4To further test the measurement properties across countries, we estimated combinations of models in which the factor loadings, correlations among constructs, and uniquenesses (i.e., error variances) were restricted to be invariant across the three samples. Comparisons of these models using practical fit indices indicated that both the factor correlations and uniqueness were noninvariant across these samples. These results are available upon request.

5The intercept can be conceptualized as the constant term in regression equations and represents a kind of baseline level for the observed variable. The intercepts estimated usually equal the actual observed mean if the model is to be perfectly reproduced. It should be noted that these equality constraints on the intercepts across the group are not of substantive interest but are required for model identification, as well as to allow meaningful interpretation of the latent mean differences.

6Bollen (1989, p. 366) indicated that a minimum condition for testing structured means in multigroup covariance structure analysis is that the factor form and factor loadings of the models be held invariant over groups. Marsh and Grayson (1994) further argued that for meaningful mean comparisons, invariance of intercepts must also be established. All testings conducted to this point suggested that these conditions were satisfied in the present investigation.

Manuscript submitted: September 25, 1995
Revision received: June 14, 1996