Running Head: VALIDITY OF THE TOPS

Confirmatory factor analysis of the Test of Performance Strategies (TOPS) among adolescent athletes

Andrew M. Lane
University of Wolverhampton, United Kingdom,

Chris Harwood
Loughborough University, United Kingdom,

Peter C. Terry
University of Southern Queensland, Australia,

and

Costas I. Karageorghis
Brunel University West London, United Kingdom

Revision submitted December 2003

Confirmatory factor analysis of the Test of Performance Strategies (TOPS) among adolescent athletes

Revision submitted December 2003

Abstract

The present study examined the factorial validity of the Test of Performance Strategies (TOPS; Thomas et al., 1999) among adolescent athletes using confirmatory factor analysis. The TOPS was designed to assess eight psychological strategies used in competition (i.e., activation, automaticity, emotional control, goal-setting, imagery, negative thinking, relaxation, and self-talk,) and eight used in practice (the same strategies except negative thinking is replaced by attentional control). National-level athletes (*N* = 584) completed the 64-item TOPS during training camps. Fit indices provided partial support for the overall measurement model for the competition items (RCFI = .92, TLI = .88, RMSEA = .05) but minimal support for the training items (RCFI = .86, TLI = .81, RMSEA = .06). For the competition items, the automaticity, goal-setting, relaxation, and self-talk scales showed good fit, whereas the activation, emotional control, imagery, and negative thinking scales did not. For the practice items, the attentional control, emotional control, goal-setting, imagery, and self-talk scales showed good fit, whereas the activation, automaticity, and relaxation scales did not. Overall, it appears that the factorial validity of the TOPS for use with adolescents is questionable at present and further development is required.

Keywords: measurement, structural equations, model testing, psychological skills

Introduction

The quest to nurture talent has contributed to greater involvement by applied sport psychologists with adolescent athletes (Bloom, 1985; Csikzentmihalyi et al., 1993; Côté, 1999). Encouraging athletes to use psychological skills during their formative years may promote better coping skills once they mature. For the practitioner and researcher alike, it might be more important to investigate the development and application of psychological skills among aspiring young athletes than among their mature counterparts. Benefits of integrating psychological skills training into youth sport programmes have been proposed by researchers in the areas of stress (e.g., Hanton and Jones 1999), achievement motivation (e.g., Harwood and Swain, 2001) and the psychological characteristics of peak performance (e.g., Gould et al., 2002), all of whom advocated developing psychological skills at a young age.

Recently, within national governing bodies of sport in the United Kingdom, increased emphasis has been placed on identifying and nurturing talented athletes via the World Class Potential plans, funded by the National Lottery®. Hence, methods of gaining insight into existing psychological skills among young athletes have an important role to play in helping to determine individual requirements for psychological skills training and practitioner support offered to athletes through such schemes.

Assessment of psychological skills is often recognised as an integral part of the work of an applied sport psychologist (see Chartrand et al., 1992; White, 1993; Smith et al., 1995; Taylor, 1995; Hardy et al., 1996; Thomas et al., 1999). One method of measuring the mental skills of athletes is by using psychological questionnaires or inventories, but the utility of such instruments depends fundamentally upon their psychometric properties. If validity and reliability have not been clearly demonstrated, it is hazardous to accept and apply data derived from such measures (Schutz and Gessaroli, 1993; Schutz, 1994).

Several researchers (see Jackson et al., 2000; Fletcher and Hanton, 2001; Gould et al., 2002) have recommended the use of the Test of Performance Strategies (TOPS; Thomas et al., 1999) as the psychometric instrument of choice for assessing psychological skills usage. The TOPS is a 64-item measure, designed to assess the "psychological processes thought to underlie successful athletic performance as delineated by contemporary theory" (Thomas et al., 1999, p. 699). Thomas and his colleagues presented a dual rationale for developing the TOPS. First, they pointed out that the validity of previous measures of psychological skills usage had not been established beyond doubt (see Murphy and Tammen, 1988). For example, the factor structure of the Psychological Skills Inventory for Sport (PSIS; Mahoney et al., 1987) was not fully supported by a subsequent validation study (Chartrand et al., 1992).

Second, they emphasised the importance of distinguishing strategies used in competition from those used during practice; a context in which many athletes spend the vast majority of their time. Thomas and associates hypothesised that eight dimensions of psychological skills – activation, attentional control, automaticity, emotional control, goal-setting, imagery, relaxation, and self-talk – would be common to both competition and practice contexts. Exploratory factor analyses supported this structure for the practice items but identified a slightly different solution for the competition items, with negative thinking replacing attentional control as a competition-specific factor.

In the development and preliminary validation of the TOPS, Thomas and colleagues recruited participants from a wide range of sports across different performance standards, from recreational to senior international. The heterogeneity of the initial validation sample is a strength of the process to validate the measure, in terms of its applicability across a wide age range and spectrum of ability. Also, in making a distinction between competition and practice strategies, the authors of the TOPS opened the way for researchers and practitioners to assess individual needs and monitor developments over time in both contexts.

However, there are at least two reasons for conducting further psychometric evaluation of the TOPS. First, Thomas and colleagues made a clear recommendation that such research should be undertaken. Specifically, they called for confirmatory factor analysis (CFA) to examine the reliability of the factor structure. Second, certain methodological features of their validation study suggest that a re-evaluation would be prudent before applying the measure to other populations of interest. For example, the ratio of participants to items was below the generally accepted criterion. Thomas and colleagues used a ratio of 4.25 (472: 111), whereas a ratio of 10:1 is usually recommended (e.g., Tabachnick and Fidell, 2001). Also, exploratory factor analysis has been criticised for producing mathematically-driven factors that are unique to the sample under investigation rather than reflective of more generalisable constructs (Schutz, 1994; Thompson and Daniel, 1996); a tendency exacerbated when the participant-to-item ratio is low (Tabachnick and Fidell, 2001). Given that Thomas et al. had proposed an *a priori* theoretical framework, it might have been more appropriate to use confirmatory techniques from the outset, either to test the complete measurement models hypothesised for the competition and practice strategies, or to first test each factor independently, with additional factors being included in subsequent analyses (see Mullan et al., 1997).

It is also interesting that, in their exploratory analysis of competition strategies, Thomas and colleagues found that items related to attentional control and emotional control formed a single factor, whereas they had originally conceptualised them as distinct constructs. They pointed out, reasonably enough, that good attentional control is inextricably linked with good emotional control. However, while attentional control and emotional control may closely co-vary they remain conceptually distinct (Lazarus, 1991, 1999) and several researchers have warned of the pitfalls of using exploratory factor analysis to generate theory (e.g., Schutz, 1994; Thompson and Daniel, 1996).

To date, no research has evaluated the validity of any measure of psychological skills usage, including the TOPS, for use among *adolescent* athletes. Consequently, it is both topical and important that researchers should address this gap in the literature. While the TOPS has shown promising signs of factorial validity, there remains a need to test the measurement model using confirmatory techniques; hence this formed the primary purpose of the present study. A secondary purpose, should the confirmatory analysis provide support for the existing measurement models, was to examine differences in adolescent athletes' usage of psychological skills in practice and competition. Based on the findings of Thomas et al. (1999), it was hypothesised that athletes would report significantly higher usage of psychological skills in competition than in practice.

Method

Participants

Participants were 584 volunteer athletes (Age range: 15-18 years, male = 264, *M* = 16.6 years, *SD* = 1.8; female = 320, *M* = 16.9 years, *SD* = 1.9). Participants were drawn from national-level training camps in the U.K., organised through a joint initiative between Nike, Inc., The Institute of Youth Sport, and the Youth Sport Trust. Athletes competed in a broad range of sports, including badminton, fencing, field hockey, lacrosse, rugby union, soccer, squash, triathlon, track and field, and volleyball.

Measurement instrument

The 64-item Test of Performance Strategies is a self-report instrument designed to measure an athlete's use of psychological skills and strategies during competition and practice (Thomas et al., 1999). Exploratory factor analysis has previously indicated an 8-factor solution for competition items and a slightly different 8-factor solution for practice items. Seven factors are common to both competition and practice contexts, whereas negative thinking is only included in the competition context and attentional control only in the practice context. Each subscale has four items. Items were rated on a 5-point scale anchored by 1 (*never*) to 5 (*always*). Scores for each subscale were summed and divided by four; resulting in overall factor scores that could range from 1 - 5.

Procedures

Written parental consent for participation in the research was granted prior to each training camp through the Youth Sport Trust. Participants also gave informed consent and volunteered to participate in the present study with no incentives. Instructions to participants included a reminder to respond to all items and a statement designed to discourage a social desirability bias (c.f., Martens et al., 1990). Participants completed the questionnaire during breaks in their training camps away from the gaze of their peers.

Model Testing

The 8-factor measurement models for practice and competition specified that items were related to their hypothesised factor with the variance of the factor fixed at 1. The first model tested allowed factors to freely intercorrelate and a second model constrained relationships between automaticity and other factors to zero. This second model was prompted by previous findings of inconsistent relationships between automaticity and other factors. For example, Thomas et al. (1999) reported minimal interrelationships whereas Thomas and Over (1994) had previously reported strong positive relationships. Given that such relationships are of theoretical and practical importance, we decided to model both possibilities.

Another approach used to evaluate the measurement model was to assess the psychometric properties of each factor independently (see Mullan et al., 1997). This approach has been proposed to be appropriate where overall model fit is not supported (Woodman and Hardy, 2003). Poor model fit can be due to various characteristics, such as when items have low loadings on their hypothesised factors, high loadings across multiple factors, or mis-specified correlations between factors. Given that researchers and practitioners will calculate factor scores by combining items in each subscale, all subscales should demonstrate factorial validity independently.

CFA using EQS V5 (Bentler, 1990, 1995; Bentler and Wu, 1995) was used to test the hypothesised models. An assumption underlying SEM is that data are normally distributed. In particular, the maximum likelihood method (ML: Chou and Bentler, 1995) used in the present study assumes multivariate normality. Two precautions were taken to guard against the effects of non-normality. The first was to utilise a large sample, above 500 participants, as recommended by Tabachnick and Fidell (2001). The second precaution, in the event of identifying multivariate non-normality, was to use the scaled X^2 Satorra-Bentler method, which has been shown to effectively control for overestimation of X^2, under-estimation of incremental fit indexes, and under-identification of errors (see West et al., 1995).

The choice of cut-off criteria used to evaluate model adequacy is a contentious issue. Some researchers favour a two-index strategy, with the indices selected on the basis of sample size, model complexity, and the distributional properties of the data (e.g., Hu and Bentler, 1999). We followed the approach of Hoyle and Panter (1995), Byrne (1998, 2000), and Kline (1998), all of whom advocated use of a range of fit indices to judge model adequacy. According to Hoyle and Panter (1995), there is little agreement among researchers about the best index of overall fit used in CFA. Consequently, to achieve a comprehensive evaluation of model fit a range of different indices were used.

Model fit was assessed using two incremental fit indices; the robust comparative fit index (RCFI: Bentler, 1995) and the non-normed fit index or Tucker-Lewis index (TLI: Tucker and Lewis, 1973). Incremental fit indices are based on comparisons between the hypothesised model and a null model (in which there are no relationships among the observed variables) and are not influenced by sample size (Marsh et al., 1988; 1996; Bentler, 1990). Kline (1998) proposed that values for the RCFI and TLI of less than .90 indicate that the hypothesised model could be substantially improved, and Hu and Bentler (1999) suggested that, in most circumstances, values should approach .95. The third index used was the root mean square error of approximation (RMSEA: Steiger, 1990), which indicates the mean discrepancy between the observed covariances and those implied by the model per degree of freedom, and therefore has the advantage of being sensitive to model complexity. A value of .05 or lower indicates a good fit and values up to .08 indicate an acceptable fit (Browne and Cudeck, 1993).

The Lagrange multiplier and Wald tests were used to highlight how model fit could be enhanced, as recommended by Biddle et al. (2001). The Lagrange test asks how can model fit be improved by adding parameters while the Wald test asks which parameters, if any, could be deleted or have their variance fixed to zero (see Tabachnick and Fidell, 2001).

Cronbach alpha coefficients were calculated to assess the internal consistency of the 16 factors. The criterion for acceptability for an internally reliable scale is normally set at .70 (Tabachnick & Fidell, 2001). However, Loewenthal (2001) suggested that an alpha of .60 is acceptable for subscales with four items.

Results

Preliminary analyses showed that multivariate normality was violated for both competition items (Mardia = 48.05) and practice items (Mardia = 49.49); indicating that use of the Satorra-Bentler estimation method was appropriate. Results of the CFA are shown in Table 1. For the competition items, fit indices showed partial support for the hypothesised measurement model. The RMSEA indicated good model fit but the RCFI fell between traditional and contemporary benchmarks; higher than the .90 criterion advocated by Kline (1998) but lower than the .95 criterion proposed by Hu and Bentler (1999). The TLI fell below the minimum value for acceptable fit, indicating that the competition model could be improved significantly. For the practice items, limited support for the model was found. The RMSEA showed support for the adequacy of the measurement model, whereas the incremental fit indices suggested that the practice model could also be improved significantly.

Given this evidence that model fit could be improved, modification indices were scrutinised to identify the specific weaknesses of the measurement models. Wald test results showed that model fit could be improved by constraining correlations between automaticity scores and other factor scores to zero. This was evident for both the competition and practice items in this factor.

The revised model had negligible impact on the fit statistics for the competition items although all fit indices for the practice items were improved (see Table 1). However, despite rectifying the model, fit indices for both competition and practice items fell short of the .95 criterion for acceptability (Hu and Bentler, 1999) and, even when judged against the less stringent traditional criterion, TLI values indicated that significant improvements to the measurement model could be made for both sets of items.

Independent analyses of each factor were then conducted. For the competition items, CFA provided strong support for the automaticity, goal-setting, relaxation, and self-talk scales, with all fit indices at the level of acceptability or better (see Table 2). Scope for improvement was indicated for the emotional control, imagery, and negative thinking scales and no support was found for the fit of the activation scale. Among the practice items, strong support was found for the attentional control, emotional control, goal-setting, imagery, and self-talk scales, scope for improvement was indicated for the automaticity and relaxation scales, and again no support was found for the activation scale.

Alpha coefficients (Cronbach, 1951) for the competition scales were activation = .69, automaticity = .74, emotional control = .72, goal setting = .80, imagery = .82, negative thinking = .75, relaxation = .80, and self-talk = .77; and for the practice scales were activation = .57, attentional control = .68, automaticity = .62, emotional control = .73, goal setting = .79, imagery = .75, relaxation = .75, and self-talk = .76. Given the 4-item scales of the TOPS, these values suggest adequate internal validity for the competition scales

(Loewenthal, 2001) and for all practice scales with the exception of activation.

Analysis at the individual item level is reported in Table 3 for the competition items and Table 4 for the practice items. The strength of an item is indicated by high factor loadings and low standard errors. Comrey and Lee (1992) suggested that factor loadings greater than .71 (50% overlapping variance) are excellent, .63 (40% overlapping variance) very good, .55 (30% overlapping variance) good, .45 (20% overlapping variance) fair, and .32 (10% overlapping variance) poor (see Tabachnick & Fidell, 2001, p. 625). By this rule of thumb, in the present analysis, 54 of the 64 items (84%) could be considered good to excellent and only 3 items (5%) could be considered poor. Of these three items, two (‘I have difficulty increasing my energy level during workouts’ and ‘I have trouble energizing myself if I feel sluggish during practice’) were in the practice activation scale and one (‘during practice, I don’t think about performance much – I just let it happen’) in the practice automaticity scale (see Table 4).

Overall, support for the TOPS measurement models was mixed for both the competition and practice scales. The scales demonstrated satisfactory psychometric properties on some tests but were shown to be in need of improvement on others. Given this partial support for the hypothesised measurement models, we decided to explore the second purpose of the study. Results of a multivariate analysis of variance are shown in Table 5. A significant multivariate effect of context was found (Wilks' $\lambda_{7,577} = .08$, $p < .001$, $\eta^2 = .66$). As hypothesised, the effect was characterised by greater use of the seven common psychological strategies by adolescent athletes in the competition environment compared to the practice environment. Univariate analyses showed these differences to be statistically significant for the strategies of activation ($\eta^2 = .53$), automaticity ($\eta^2 = .13$), emotional control ($\eta^2 = .07$), goal setting ($\eta^2 = .09$), imagery ($\eta^2 = .11$), and relaxation ($\eta^2 = .53$). However, considering the psychometric properties of the activation scales, the observed difference in use of this strategy in the competition and practice situations should be interpreted with great caution. Indeed, caution is advised in the interpretation of between-context differences for *all* TOPS subscales.

Discussion

The purpose of the present study was to investigate the factorial validity of the TOPS for use with adolescent athletes and, subsequently, to advance awareness of psychological skills usage in this population. The TOPS assesses the use of psychological strategies in competition and practice settings. Such measures serve several important functions. First, the TOPS can be used as a research tool into the effectiveness of interventions to improve psychological skills. Second, TOPS scores provide information for coaches and practitioners about the existing psychological skills and future needs of young athletes. Third, information derived from the TOPS can help to educate governing bodies about those psychological skills and qualities that are being properly addressed or perhaps neglected during the early stages of player development.

However, the utility of measures for assessing psychological skills depends upon the demonstration of appropriate validity characteristics. Although Thomas et al. (1999) provided some evidence to support the factorial validity of the TOPS among a heterogeneous sample of adults and adolescents, researchers should test the generalisability of measures for different populations of interest (Schutz and Gessaroli, 1993; Schutz, 1994; Anastasi and Urbina, 1997). Researchers should also use the most stringent methods to assess validity (Schutz, 1994). The present study was designed to address both objectives by testing the measurement models for competition and practice presented by Thomas and associates, using confirmatory techniques.

Results indicated that neither measurement model adequately fitted the data although the competition model showed better fit than the practice model. At the subscale level, many scales showed very good fit, other less so. More specifically, for the competition items, the automaticity, goal-setting, relaxation, and self-talk scales showed good fit, whereas the activation, emotional control, imagery, and negative thinking scales could be improved significantly. For the practice items, the attentional control, emotional control, goal-setting, imagery, and self-talk scales showed good fit, whereas the activation, automaticity, and relaxation scales could be improved.

At the individual item level, 84% of items showed good to excellent loadings on their hypothesised factor. Some specific items showed weak loadings, in particular on the practice scales for activation and automaticity. Several possible steps could be taken to address these limitations. For example, weak factor loadings can indicate that participants did not comprehend the meaning of an item in the context of the factor it was intended to represent. In such cases, an item might be reworded to aid clarity. Also, given that the original validation procedures were geared primarily towards adult athletes, it is possible that the language used in some items is inappropriate for adolescents. For example, in the automaticity scale, the item "during competition I perform on *automatic pilot*" requires understanding of the concept of an automatic navigation system to give a meaningful response. Similarly, another automaticity item, "during practice, I don't think about performing much – I just let it happen", could be interpreted by some participants as relating to performing with low motivation rather than performing with minimal processing.

Another area where changes might improve model fit involves separating the constructs of attentional and emotional control among the competition items. For example, the item, "when I make a mistake in competition, I have trouble getting my concentration back on track", shows a modest loading (.41) on emotional control, its hypothesized factor. It is difficult to support the notion that this item truly belongs as part of the *emotional* control factor when it is so clearly oriented towards *attentional* control. It is suggested that the decision by Thomas and colleagues to combine these two constructs into a single scale should be reconsidered, at least in the context of adolescent athletes.

Collectively, confirmatory techniques showed the factorial structure of the TOPS to be strong in parts but in need of revision elsewhere. In particular, both activation scales appear to require revision. Such revisions should be completed before the TOPS can be applied with confidence among adolescent populations.

In terms of psychological skills usage, the results showed that adolescent athletes used the strategies of interest more in competition settings than in the practice environment. This finding is consistent with the results previously reported by Thomas and colleagues. Future research might explore the perceived *effectiveness* of using psychological skills in these different settings rather than simply comparing the relative usage in the two domains.

It has been suggested previously (Thomas et al., 1999) that the TOPS may be an appropriate measure for investigating the extent to which psychological skills usage varies over time. Indeed, practitioners may already be using changes in TOPS scores as an index for evaluating the impact of psychological skills training. Such a strategy may be premature until the stability of the TOPS factor structure is established. Unless it can be demonstrated that scores on a questionnaire are stable under conditions where no change is expected, any changes observed may be caused by random error potentially associated with participants not understanding the meaning of items (see Nevill et al., 2001). Therefore, before meaningful comparisons can be made on TOPS data collected over time, it is important for future research to be conducted to investigate the stability of the measure.

Conclusions

In conclusion, findings of the present study have provided evidence of the psychometric properties of the TOPS among adolescent athletes. Results have shown that, in its present form, question marks remain over some aspects of the factorial validity of the measure and hence its appropriateness for use with adolescent athletes remains in doubt.

Author note

Andrew Lane, School of Sport, Performing Arts and Leisure, University of Wolverhampton, School of Sport Performing Arts and Leisure, Walsall Campus, Gorway Road, Walsall, WS1 3BD, U.K. Email: A.M.Lane2@wlv.ac.uk Tel: +44 (0)1902 322862 Fax: +44 (0)1902 323230. Peter C. Terry, Department of Psychology, University of Southern Queensland, Australia. Chris Harwood, Sport and Exercise Psychology Research Group, Dept. of Physical Education, Sports Science and Recreation Management, Loughborough University, U.K. Costas Karageorghis, Department of Sport Sciences, Brunel University, U.K.

The authors would like to thank Sam Barney for his helpful comments during the production of this paper.

References

Anastasi, A. and Urbina, S. (1997). *Psychological testing* (7th ed.). New York: Prentice-Hall.

Bentler, P.M. (1990). Comparative fit indices in structural models. *Psychological Bulletin,* **107**, 238-246.

Bentler, P.M. (1995). *EQS structural equation program manual.* Encino, CA: Multivariate Software.

Bentler, P.M. and Wu, E.J.C. (1995). *EQS/Windows user's guide.* Encino, CA: Multivariate Software.

Biddle, S.J.H., Markland, D., Gilbourne, D., Chatzisarantis, N. and Sparkes, A.C. (2001). Research methods in sport and exercise psychology: Quantitative and qualitative issues. *Journal of Sports Sciences,* **19**, 777-809.

Bloom, B.S. (1985). *Developing talent in young people*. New York: Ballatine.

Browne, M.W. and Cudeck, R. (1993). Alternative ways of assessing model fit. In *Testing structural equation models* (edited by K.A. Bollen and J.S. Long), pp. 132-162. Newbury, CA: Sage.

Byrne, B.M. (1998). *Structural equation modelling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications and programming.* Mahwah, NJ: Lawrence Erlbaum.

Byrne, B.M. (2000). *Structural equation modelling with AMOS: Basic concepts, applications and programming.* Mahwah, NJ: Lawrence Erlbaum.

Chartrand, J.M., Jowdy, D.P. and Danish, S.J. (1992). The Psychological Skills Inventory for Sports: Psychometric characteristics and applied implications. *Journal of Sport and Exercise Psychology,* **14**, 405-413.

Chou, C. and Bentler, P. (1995). Estimates and tests in structural equation modelling. In *Structural equation modelling: Concepts, issues, and applications* (edited by R.H. Hoyle), pp. 158-175. London: Sage.

Comrey, A.L. and Lee, H.B. (1992). *A first course in factor analysis* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Côté, J. (1999). The influence of the family in the development of talent in sport. *The Sport Psychologist*, **13**, 395-417.

Cronbach, L.J. (1951). Coefficient alpha and internal structure of tests. *Psychometrika,* **16,** 297-334.

Csikzentmihalyi, M., Rathunde, K., Whalen, S. and Wong, M. (1993). *Talented teenagers: The roots of success and failure*. New York: Cambridge University Press.

Fletcher, D. and Hanton, S. (2001). The relationship between psychological skills usage and competitive anxiety responses. *Psychology of Sport and Exercise,* **2**, 89-101.

Gould, D., Dieffenbach, K. and Moffatt, A. (2002). Psychological characteristics and their development in Olympic champions. *Journal of Applied Sport Psychology,* **14**, 172-204.

Hanton, S. and Jones, J.G. (1999). The acquisition and development of cognitive skills and strategies: I. Making the butterflies fly in formation. *The Sport Psychologist,* **13,** 1-21.

Hardy, L., Jones, J.G. and Gould, D. (1996). *Understanding psychological preparation for sport: Theory and practice of elite performers*. Chichester, UK: Wiley.

Harwood, C. and Swain, A. (2001). The development and activation of achievement goals in tennis: I. Understanding the underlying factors. *The Sport Psychologist,* **15***,* 319-341.

Hoyle, R.H. and Panter, A. T. (1995). Writing about structural equation models. In *Structural equation modelling: Concepts, issues, and applications* (edited by R.H. Hoyle), pp. 158-175. London: Sage.

Hu, L. and Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling,* **6***,* 1-55.

Jackson, S.A., Thomas, P.R., Marsh, H.W. and Smethurst, C.J. (2000). Relationships between flow, self-concept, psychological skills, and performance. *Journal of Applied*

Sport Psychology, **13**, 129-153.

Kline, P. (1998). *Handbook of psychological testing*. London: Routledge.

Lazarus, R.S. (1991). *Emotion and adaptation.* New York: Oxford University Press.

Lazarus, R.S. (1999). The cognition-emotion debate: A bit of history. In *Handbook of cognition and emotion* (edited by T. Dalgleish and M. Power), pp. 3-19. New York: Wiley.

Loewenthal, K.M. (2001). *An introduction to psychological tests and scales* (2nd ed.). London: UCL Press.

Mahoney, M.J., Gabriel, T.J., and Perkins, T.S. (1987). Psychological skills and exceptional athletic performance. *The Sport Psychologist,* **1**, 181-199.

Marsh, H.W., Balla, J.R. and Hau, K.T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical properties. In *Advanced structural equation modelling* (edited by G.A. Marcoulides and R.E. Schumaker), pp. 315-353. Mahwah, NJ: Lawrence Erlbaum.

Marsh, H.W., Balla, J.R. and McDonald, R.P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin,* **103**, 391-410.

Martens, R., Vealey, R.S. and Burton, D. (1990). *Competitive anxiety in sport*. Champaign, IL: Human Kinetics.

Mullan, E., Markland, D. and Ingledew, D.K. (1997). A graded conceptualization of self-determination in the regulation of exercise behavior: Development of a measure using confirmatory factor analysis procedures. *Personality and Individual Differences*, **23**, 745–752.

Murphy, S.M. and Tammen, V. (1988). In search of psychological skills. In *Advances in sport and exercise psychology measurement* (edited by J.L.Duda), pp. 195- 209. Morgantown. WV: Fitness Information Technology.

Nevill, A., Lane, A.M., Kilgour, L., Bowes, N. and Whyte, G. (2001). Stability of psychometric questionnaires. *Journal of Sports Sciences,* **19,** 273-278.

Schutz, R.W. (1994). Methodological issues and measurement problems in sport psychology. In *International perspectives on sport and exercise psychology* (edited by S.Serpa, J.Alves and V.Pataco), pp. 35-57. Morgantown, VA: Fitness Information Technology.

Schutz, R.W. and Gessaroli, M.E. (1993). Use, misuse, and disuse of statistics in psychology research. In *Handbook of research on sport psychology* (edited by R.N.Singer, M.Murphey and L.K.Tennant), pp. 901-921. New York: McMillan.

Smith, R.E., Schutz, R.W., Smoll, F.L. and Ptacek, J.T. (1995) Development and validation of a multidimensional measure of sport-specific psychological skills: The Athletic Coping Skills Inventory-28. *Journal of Sport and Exercise Psychology,* **17**, 379-398.

Steiger, J.H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research,* **25**, 173-180.

Tabachnick, B.G. and Fidell, L.S. (2001). *Using multivariate statistics* (4th ed.). New York: Harper and Row.

Taylor, J. (1995). A conceptual model for integrating athletes' needs and sport demands of competitive mental preparation strategies. *The Sport Psychologist,* **9**, 339-357.

Thomas, P.R. and Over, R. (1994). Psychological and psychomotor skills associated with golf performance. *The Sport Psychologist,* **8**, 73-86.

Thomas, P.R., Murphy, S. and Hardy, L. (1999). Test of Performance Strategies: Development and preliminary validation of a comprehensive measure of athletes' psychological skills. *Journal of Sports Sciences,* **17**, 697-711.

Thompson, B. and Daniel, L.G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological*

Measurement, **56***,* 197-208.

Tucker, L.R. and Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika,* **38**, 1-10.

West, S.G., Finch, J.F. and Curren, P.J. (1995). Structural equations models with nonnormal variables: Problems and remedies. In *Structural equation modelling: Concepts, issues, and applications* (edited by R.H.Hoyle), pp. 56-73. London: Sage.

White, S.A. (1993). The relationship between psychological skills, experience, and practice commitment among collegiate male and female skiers. *The Sport Psychologist,* **7**, 49-57.

Woodman, T. and Hardy, L. (2003). The relative impact of cognitive anxiety and self-confidence upon sport performance: A meta-analysis. *Journal of Sports Sciences,* **21***,* 443-457.

Table 1

Confirmatory factor analysis of the Test of Performance Strategies

Fit index	8-factor model for Competition	8-factor model for Practice	Revised model for Competition	Revised model for Practice
χ^2	901	1079	925	801
df	436	436	443	360
TLI	.88	.81	.88	.86
RCFI	.92	.86	.91	.90
RMSEA	.05	.06	.05	.05

Note. All χ^2 values are significant at $p < .01$.

Table 2

Confirmatory factor analysis of subscales of the Test of Performance Strategies

Subscales	χ^2	TLI	RCFI	RMSEA
Competition				
Activation	40.47	.68	.88	.19
Automaticity	4.24	.98	.99	.06
Emotional control	29.54	.78	.93	.18
Goal-setting	1.53	1.00	1.00	.00
Imagery	22.23	.90	.98	.16
Negative thinking	29.54	.78	.93	.18
Relaxation	2.02	1.00	1.00	.03
Self-talk	.65	1.00	1.00	.00
Practice				
Activation	58.57	.29	.73	.23
Attentional control	3.55	.97	.99	.06
Automaticity	11.05	.89	.96	.01
Emotional control	3.74	.98	1.00	.06
Goal-setting	3.21	.99	1.00	.04
Imagery	.39	1.00	1.00	.00
Relaxation	23.50	.87	.95	.15
Self-talk	4.11	.98	1.00	.05

Table 3

Factor loadings for competition items by subscale

Subscale	Item	Factor loading	Error
Goal-setting			
	I set very specific goals for competition	.752	.659
	I set personal performance goals for a competition	.751	.661
	During competition I set specific result goals for myself	.730	.683
	I evaluate whether I achieve my competition goals	.595	.804
Automaticity			
	During competition, I don't think about performing much - I just let it happen	.732	.682
	During competition I perform on 'automatic pilot'	.660	.751
	During competition, I play/perform instinctively with little conscious effort	.650	.760
	I perform at competitions without consciously thinking about it	.549	.836
Emotional control			
	My emotions keep me from performing my best at competitions	.811	.585
	My emotions get out of control under the pressure of competition	.776	.631
	When something upsets me during a competition, my performance suffers	.469	.883
	When I make a mistake in competition, I have trouble getting my concentration back on track	.412	.911
Imagery			
	I rehearse my performance in my mind at competitions	.845	.535
	I imagine my competitive routine before I do it at a competition	.815	.579
	At competitions, I rehearse the feel of my performance in my imagination	.710	.704
	I visualize my competition going exactly the way I want	.559	.829
Activation			
	I do what needs to be done to get psyched up for competitions	.740	.672
	I psych myself up at competitions to get ready to perform	.599	.801
	I can increase my energy to just the right level for competitions	.573	.819
	I can raise my energy level at competitions when necessary	.490	.872
Self-talk			
	I say things to myself to help my competitive	.736	.677

	performance		
	I manage my self-talk effectively during competition	.716	.698
	I have specific cuewords or phrases that I say to myself to help my performance during competition	.541	.841
	I talk positively to myself to get the most out of competitions	.759	.651
Relaxation			
	I am able to relax if I get too nervous at a competition	.815	.580
	When I need to, I can relax myself at competitions to get ready to perform	.745	.667
	When pressure is on at competitions, I know how to relax	.708	.706
	I find it difficult to relax when I am too tense at competitions	.571	.821
Negative thinking			
	My self-talk during competition is negative	.688	.725
	I keep my thoughts positive during competitions	.648	.761
	During competition I have thoughts of failure	.644	.765
	I imagine screwing up during a competition	.634	.773

Table 4

Factor loadings for practice items by subscale

Subscale	Item	Factor loading	Error
Goal setting			
	I have very specific goals for practice	.861	.508
	I set goals to help me use practice time effectively	.696	.718
	I don't set goals for practices, I just go out and do it	.620	.785
	I set realistic but challenging goals for practice	.644	.765
	Competition		
Automatcity			
	During practice, my movements and skills just seem to flow naturally from one to another	.678	.735
	During practice sessions I just seem to be in a flow	.624	.782
	At practice, I can allow the whole skill or movement to happen naturally without concentrating on each part of the skill	.583	.813
	During practice, I don't think about performing much - I just let it happen	.323	.947
Emotional control			
	I have trouble controlling my emotions when things are not going well at practice	.838	.546
	When things are going poorly in practice, I stay in control of myself emotionally	.602	.799
	When I perform poorly in practice I lose my focus	.587	.809
	I get frustrated and emotionally upset when practice does not go well	.551	.835
Imagery			
	At practice, when I visualize my performance, I imagine what it will feel like.	.729	.684
	I rehearse my performance in my mind before practice	.681	.732
	At practice, when I visualize my performance, I imagine watching myself as if on a video replay	.615	.788
	During practice I visualize successful past performances	.605	.796
Activation			
	I practise energizing myself during training sessions	.724	.690
	I practise a way to energize myself	.658	.753
	I have difficulty increasing my energy level during workouts	.286	.958

	I have trouble energizing myself if I feel sluggish during practice	.264	.965
Self-talk			
	I motivate myself to train through positive self-talk	.754	.657
	I talk positively to myself to get the most out of practice	.696	.718
	I say things to myself to help my practice performance	.614	.789
	I manage my self-talk effectively during practice	.596	.803
Relaxation			
	I practise using relaxation techniques at workouts	.809	.588
	I practise a way to relax	.778	.628
	I use practice time to work on my relaxation technique	.650	.760
	I relax myself at practice to get ready	.415	.910
Attentional control			
	During practice I focus my attention effectively	.664	.748
	My attention wanders while I am training	.633	.774
	I have trouble maintaining my concentration during long practices	.579	.815
	I am able to control distracting thoughts when I am training	.512	.859

Table 5

Comparison of competition and practice scores on the Test of Performance Strategies among 584 adolescent athletes

	Competition		Practice			
	M	*SD*	*M*	*SD*	*F*	η^2
Activation	3.73	.70	2.98	.64	654.11*	.53
Automaticity	3.73	.70	2.81	.80	89.19*	.13
Emotional control	3.42	.75	3.23	.78	46.42*	.07
Goal-setting	3.49	.84	3.27	.82	58.53*	.09
Imagery	3.15	.93	2.93	.89	70.12*	.11
Relaxation	3.21	.81	2.17	.76	663.92*	.53
Self-talk	3.29	.88	3.26	.81	1.94	.00
Attentional control			3.47	.64		
Negative thinking	2.42	.78				

Wilks' $\lambda_{7,577} = .34$, $p < .001$, $\eta^2 = .66$

* $p < .01$