Psychology of Sport and Exercise 11 (2010) 27–35

Contents lists available at ScienceDirect

Psychology of Sport and Exercise

journal homepage: www.elsevier.com/locate/psychsport

Test of Performance Strategies (TOPS): Instrument refinement using confirmatory factor analysis

Lew Hardy [a,*], Ross Roberts [a], Patrick R. Thomas [b], Shane M. Murphy [c]

[a] Institute for the Psychology of Elite Performance, School of Sport, Health and Exercise Science, George Building, Bangor University, Holyhead Road, Bangor, Gwynedd LL57 2PZ, United Kingdom
[b] School of Education and Professional Studies, Mt Gravatt Campus, Griffith University, Queensland QLD 4111, Australia
[c] Department of Psychology, Western Connecticut State University, 181 White Street, Danbury CT 06810, USA

ARTICLE INFO

Article history:
Received 24 November 2008
Received in revised form
2 March 2009
Accepted 23 April 2009
Available online 3 May 2009

Keywords:
Sport psychology
Questionnaire
Psychological skills
Practice
Competition

ABSTRACT

Background and purpose: A recent confirmatory factor analysis (CFA) of the Test of Performance Strategies (TOPS) by Lane, Harwood, Terry, and Karageorghis [2004. Confirmatory factor analysis of the Test of Performance Strategies (TOPS) among adolescent athletes. *Journal of Sports Sciences, 22*, 803–812] provided only mixed support for structural integrity of the TOPS. The objectives of the present paper were to further examine the instrument's structural integrity and enhance it if necessary.

Method and results: In a pilot study, a sample of North American athletes completed the TOPS. Results revealed poor fits during analysis of the competition and practice subscales. In Study 1, a number of new items were developed and a new competition subscale (distractibility) introduced, to address the problems identified and create the TOPS 2. CFAs of responses from a sample of Australian, North American and British athletes provided much stronger support for the factorial validity of the TOPS 2 inventory. However, the distractibility subscale suffered from poor factor loadings and reliability, and so was removed from further analysis. In Study 2, the factorial validity of the TOPS 2 was confirmed on a new sample of Australian athletes.

Conclusions: The TOPS 2 appears to be an improvement over the TOPS. Implications of the results for practitioners are discussed, and future research directions are recommended.

Psychological inventories, based on athlete self-report, are an important means of assessing the cognitive and affective states of athletes (Vealey & Garner-Holman, 1998). Indeed, some argue that psychological assessment is an essential requirement for any sport psychology intervention (e.g., Beckmann & Kellmann, 2003). Traditional forms of assessment in sport psychology were based on the assumption that personality traits or states determine an individual's pattern of behaviour (Tkachuk, Leslie-Toogood, & Martin, 2003). Thus, sport psychologists used some instruments from clinical and counselling psychology, such as the 16 Personality Factor Questionnaire (Cattell, 1949), the State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970), and the Profile of Mood States (McNair, Lorr, & Droppleman, 1971). Other instruments were developed specifically to measure such constructs with athletes, including the Athletic Motivation Inventory (Tutko, Lyon, & Ogilvie, 1969), the Test of Attentional and Interpersonal Style (Nideffer, 1976), and the Sport Competition Anxiety Test (Martens, 1977). However, the suitability of these measures in sport psychology research and practice has been questioned because of their clinical focus, absence of athlete norms, questionable psychometric properties, and inconsistencies in findings (Ford & Summers, 1992; LeUnes & Nation, 1989; Tkachuk et al., 2003).

Many sport psychologists have turned to inventories that measure sport-related behaviours rather than any underlying personality dimensions that might be linked to those behaviours. There is considerable interest in instruments targeting psychological skills and strategies in sport as they are likely to differentiate more and less successful athletes, and provide evidence regarding the efficacy of psychological skills training programs. Such instruments include the Psychological Skills Inventory for Sport (PSIS, Mahoney, Gabriel, & Perkins, 1987); the Athletic Coping Skills Inventory (ACSI-28, Smith, Schutz, Smoll, & Ptacek, 1995); the Ottawa Mental Skills Assessment Tool (OMSAT-3, Durand-Bush, Salmela, & Green-Demers, 2001); and the Test of Performance Strategies (TOPS, Thomas, Murphy, & Hardy, 1999). Such instruments must have sound psychometric properties to be of any theoretical or practical value.

* Corresponding author. Fax: +44 1248 371053.
E-mail address: l.hardy@bangor.ac.uk (L. Hardy).

1469-0292/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.psychsport.2009.04.007

The TOPS was designed to measure a comprehensive range of psychological skills and techniques, and their strategic use by athletes both in competition and at practice. Subscales were developed targeting eight of the most salient psychological skills and processes thought to underlie successful athletic performance. Interested readers are referred to Thomas et al. (1999) for more on the theoretical basis for the inclusion of the particular psychological skills within the TOPS scales. These skills are goal setting, relaxation, activation, imagery, self-talk, attentional control, emotional control, and automaticity. All of these skills and strategies are measured at practice, but as a result of exploratory factor analyses by Thomas et al., negative thinking rather than attentional control is measured in competition. Responses from the original sample of athletes revealed moderate correlations among many of the TOPS subscales – automaticity being a notable exception, as well as similarities in the patterns of use of these skills and strategies at practice and in competition (Thomas et al., 1999). Those performing at international and national standards reported more use of most psychological skills and strategies and less negative thinking than club/recreational athletes (Thomas et al., 1999). Automaticity was again an exception. Older athletes reported more automaticity in the execution of skills than younger athletes. However, it is not clear why male athletes performing at international and national standard reported lower levels of automaticity in competition than less skilled male athletes. There was no clear pattern of differences in female athletes.

Subsequent research has provided further evidence of the construct validity and internal consistency of the TOPS subscales. Gould, Dieffenbach, and Moffett (2002) included the TOPS in a battery of inventories used to examine the psychological characteristics of Olympic champions. These elite athletes were high on goal setting, activation, relaxation, and emotional control in competition, and on goal setting and attentional control at practice. Compared to the norms for international athletes (Thomas et al., 1999), the Olympic champions scored substantially higher on automaticity, emotional control, and relaxation in competition, and lower on negative thinking. In addition, they scored substantially higher on goal setting and attentional control at practice. Gould et al. (2002) also interviewed the Olympic champions, as well as a coach, and a parent, sibling, or significant other for each of the athletes. The themes emerging from these interviews provided triangulation for many of the quantitative findings in this study, which in turn corroborated previous research identifying mental skills as key components of peak performance (Williams & Krane, 2001).

Jackson, Thomas, Marsh, and Smethurst (2001) investigated relationships between the use of psychological skills in competition, athlete self-concept, the flow experience, and performance. They found that the more athletes use psychological skills in their sport, the more likely they are to experience flow. All of the TOPS subscales, with the surprising exception of automaticity, contributed significantly to the predictions of flow. Further, there were strong correlations across instruments for measures of similar constructs (e.g., goal setting on the TOPS was correlated with the clear goals dimension of the Flow State Scale) and no correlations between TOPS subscales and dissimilar constructs (e.g., transformation of time). Internal consistency of the TOPS competition subscales was generally higher than in the original study (alphas ranged from 0.77 to 0.88).

Fletcher and Hanton (2001) used the TOPS to examine whether the use of psychological skills was related to athletes' competitive anxiety responses. They found that athletes who made high use of relaxation, self-talk, and imagery skills differed significantly in their competitive anxiety responses from those who made low use of these psychological skills. Those who made high use of relaxation reported lower levels of cognitive and somatic anxiety, which they interpreted as facilitative rather than debilitative. They also reported higher levels of self-confidence than those who made low use of relaxation. Those making high use of self-talk and imagery were more self-confident than those making low use of these skills. No differences in competitive anxiety responses were found for goal-setting usage.

The TOPS responses of a small sample of male soccer players competing at the 2000 Amputee World Cup were used to examine relationships between psychological skills, self-efficacy, and performance (Lowther, Lane, & Lane, 2002). Higher use of relaxation strategies in competition was associated with higher levels of self-efficacy, and the use of imagery at training was significantly related to self-ratings of performance in competition.

The importance of using mental skills at practice, not just in competition, is also apparent in other research with the TOPS. Frey, Laguna, and Ravizza (2003) demonstrated that athletes' use of mental skills was related to their perceived success at practice and in competition. The more athletes used mental skills at practice, the more successful they perceived themselves to be, not only at practice, but also in competition. Frey et al. (2003) urged sport psychology consultants to make coaches aware of the relationship between mental skill use in practice and success in competition, so that athletes would use mental skills when practising their physical skills, thereby enhancing the quality of practice.

The TOPS has become one of the most popular tests in sport psychology (Weinberg & Gould, 2003), with very promising subscales (Bond & Sargent, 2004). The preliminary analyses reported by Thomas et al. (1999) showed that the subscales had good construct validity, which is also evident from subsequent research. Some of the findings in relation to automaticity, however, have been contrary to expectations, suggesting that further attention should be given to the items on this subscale. Most of the subscales have demonstrated good levels of internal consistency, particularly given that they are each comprised of just four items. However, two of the subscales, measuring activation and automaticity at practice, displayed only moderate reliability in the original study (Cronbach's alphas of 0.66 and 0.67 respectively), again suggesting that these subscales warrant further attention.

Having said all this, it must be noted that the TOPS subscales were based on exploratory factor analyses. Although it is often necessary to use exploratory factor analysis as a preliminary step in inventory development, this should be followed by confirmatory factor analysis (CFA) to test the proposed factor structure (Schutz & Gessaroli, 1993). Moreover, the CFA should be conducted on data from an independent sample to cross-validate the factor structure (Schutz & Gessaroli, 1993). Thomas et al. (1999) thus signalled the need for further research using CFA to examine the factorial validity of the TOPS.

Lane, Harwood, Terry, and Karageorghis (2004) recently undertook such a confirmatory approach, and examined the factorial validity of the TOPS with data from adolescent British athletes. There were mixed results regarding the hypothesised measurement models for competition and practice. The root mean square error of approximation (RMSEA) indicated good model fit for competition, and adequate fit for practice. However, incremental fit indices (the robust comparative fit index, RCFI; and the Tucker-Lewis index, TLI) indicated that both models could be significantly improved. The analyses provided strong support for the automaticity, goal setting, relaxation, and self-talk competition subscales; indicated scope for improvement in the emotional control, imagery, and negative thinking subscales; and provided no support for the fit of the activation subscale. Corresponding analyses for the practice subscales provided strong support for the attentional control, emotional control, goal setting, imagery, and self-talk subscales; indicated scope for improvement in the automaticity and relaxation subscales; and again provided no support for the activation subscale. All competition subscales displayed adequate internal consistency (alphas ranged from 0.69 to 0.82), but two of the practice subscales, activation and automaticity, had low internal consistency (alphas of 0.57 and 0.62 respectively). Most of the items (84%) had good to excellent factor loadings and standard errors. Only 3 of the 64 items were considered poor – two on the activation in practice subscale, and one on the automaticity in practice subscale.

The participants in the Lane et al. (2004) study were all younger than 18 years of age, and the authors questioned whether the language used in some items might be inappropriate for adolescents. For example, they questioned whether adolescents understand the concept of performing "on automatic pilot", even though the CFA provided strong support for the automaticity competition subscale. Clearly, the factor structure of the TOPS needs to be further examined using data from athletes ranging in age and ability. In addition, subtle cultural differences can affect instrumentation (Gauvin & Russell, 1993). In the present paper, we therefore briefly report a pilot study which used CFA to examine the factorial validity of the TOPS for North American rather than Australian athletes from whom the original data were obtained. Our first study then examined the factorial validity of a revised version of the TOPS (TOPS 2) which contained several new items, and a new subscale developed to address measurement issues identified with the initial instrument (these issues are discussed in detail in the Method section of Study 1 under *Instrument Refinement*). The participants were a mixed sample of American, Australian, and British athletes. Finally a second study used a sample of Australian athletes to confirm the factorial validity of the TOPS 2.

Pilot study

A sample ($n = 520$) of North American male (M age = 16.97 years; $SD = 1.99$) and female athletes (M age = 17.54 years; $SD = 2.77$) of varying ability levels completed the TOPS (Thomas et al., 1999) away from their competition and training environments in small groups of 5–20. All athletes gave their informed consent to complete the TOPS and parental consent was provided for those athletes under the age of 18. CFA, utilising a sequential model testing approach that tests models at a single factor level, pairwise level and full model level (see Biddle, Markland, Gilbourne, Chatzisarantis, & Sparkes, 2001; Jöreskog, 1993; and Study 1 of the present manuscript), was employed to analyse the practice and competition subscales separately.

Results

Competition subscales

At the single factor and pairwise levels, problems were found for the emotional control, negative thinking, and activation subscales. Thus, due to the number of problematic subscales within the competition inventory it was felt that there were insufficient grounds with which to analyse the full eight-factor competition model.

Practice subscales

The single factor models highlighted activation as a problematic subscale. Furthermore, the pairwise models highlighted that automaticity, activation, and attentional control were problematic. In light of the number of problematic practice subscales, the full eight-factor practice model was not analysed.

Discussion

The analyses revealed a number of problematic subscales with the TOPS. Furthermore, some of these factors (e.g., activation) had been previously identified as problematic (Lane et al., 2004; Thomas et al., 1999). In light of the problematic subscales, the authors felt that there was clearly a need to develop further items on the problem subscales and then re-evaluate the fit of the modified inventory.

Study 1

Method

Participants

The participants were 220 Australian, 120 North American, and 225 British male and female athletes of varying ability levels from 48 different sports. Participants completed the questionnaire away from their competition and training environments in small groups of 5–20 participants. Responses indicated that the mean age of the Australian males was 24.1 years ($SD = 11.96$ years). The mean age of the Australian females was 20.13 years ($SD = 9.33$ years). The mean age for the North American males was 17.2 years ($SD = 1.90$ years); for North American females, it was 14.74 years ($SD = 2.1$ years). The mean age for the British males was 21.71 years ($SD = 5.85$ years), and the mean age for the British females was 20.13 years ($SD = 2.92$ years).

Participants' responses also indicated that 9.50% of the Australian participants, 22.50% of the North American participants, and 3.60% of the British participants were competing internationally. Of the Australian participants, 21.80% were competing nationally, as were 74.20% of the North American participants and 8% of the British participants. Of the rest, 30.50% of the Australian participants, 0.8% of the North American participants, and 14.20% of the British participants were competing at a regional or junior national level; the remainder were competing at a club level. All participants gave their informed consent to participate in the study.

Instrument refinement

The results of the pilot study, together with the results of Lane et al.'s (2004) study, indicated a clear need to re-examine the items that had been used to develop the activation and emotional control subscales in both the practice and the competition domains. The original conceptualisation of activation by Thomas et al. (1999) had been in terms of readiness to perform (cf. Hardy, Jones, & Gould, 1996; Pribram & McGuiness, 1975) rather than unidimensional arousal levels, and it was felt that items retained following the original factor analysis did not fully reflect this. Consequently, two of the original competition activation items were retained and a further four were developed. Similarly, one of the original practice activation items was retained and a further five were developed. All of these new items reflected readiness to perform.

In the original factor analysis of the TOPS, it proved impossible to isolate an attentional control in competition factor. The items that had been generated to assess this dimension migrated either to the emotional control factor (with negative loadings) or to a factor that had not been predicted by the authors, but which was subsequently labelled negative thinking. Consequently, for the present study further items were generated in order to purify the emotional control in competition subscale. Three of the original items were retained and three new items were generated. Although the emotional control in practice subscale had not been problematic, the opportunity was taken to examine a further two items in this subscale so that parallel practice and competition subscales could be produced. Because the positively phrased attentional control in competition items had all loaded ambiguously, while the negatively phrased items had loaded on negative thinking and emotional control, the authors considered the possibility that a resistance to disruption/distractibility factor might be a more appropriate way of measuring attentional control in competition. Consequently, one of the original attentional control items (reflecting distractibility) was retained and a further 10 items reflecting either distractibility or resistance to disruption were generated.

Examination of the relaxation in practice subscale raised an interesting issue. In the original TOPS, relaxation in practice contained three items that clearly assessed the strategic use of relaxation and one item that referred to practising a relaxation strategy. This item could be construed to confound skill level with strategic use. Conversely, the relaxation in competition subscale contained four items that referred to the level of skill the respondent had in terms of utilising relaxation in competition. However, all of the items in the original TOPS subscales assessing the other three basic psychological skills of goal setting, imagery, and self-talk (cf. Hardy et al., 1996; Vealey, 1988) refer to strategic use of these skills in both the competition and practice arenas. Furthermore, both Vealey and Hardy and associates present arguments that performers who routinely practice using basic psychological skills will eventually demonstrate higher levels of skill at the more advanced psychological skills (e.g., emotion and attention control). Consequently, a decision was made to modify the relaxation items so that all the subscales measuring Vealey's four basic psychological skills (goal setting, imagery, relaxation, and self-talk) were phrased with reference to the strategic use of those skills, whilst all the subscales measuring more advanced psychological skills were phrased with reference to skill levels. This decision to modify the relaxation subscale led to the retention of three of the practice items and the generation of three new items for that domain, plus the generation of six new items for the competition domain.

Finally, a decision was made to modify the automaticity items for both the practice and competition domains. As well as the psychometric problems that had been shown to exist with the practice subscale, the authors were concerned some of the items in the original TOPS might lead respondents to confuse automaticity with a laissez faire attitude. Consequently, for the practice subscale, five new items were generated which better reflected the target construct, with one of the original TOPS items being retained and several others being modified. For the competition subscale, six new items were generated. This new version of the Test of Performance Strategies (TOPS 2) thus initially consisted of the 64 items from the original TOPS together with 49 new items. The new items were developed by the first and third authors following extensive theoretical reviews of the constructs under consideration and searches of existing measures that might be relevant. Their clarity and face validity was confirmed by two other experienced sport psychologists. Where either sport psychologist identified a problem, the item was revised.

Procedure

The TOPS 2 was administered to participants either individually or in small groups of 10–20 people. Participants received no financial or other tangible reward for their participation, but did receive a copy of their own TOPS 1 profile together with the table of norms produced by Thomas and associates. No other measures were administered with the TOPS 2.

Analytical strategy

In line with Jöreskog's (1993) recommendations, the factorial validity of the TOPS 2 was examined via analysis of covariance structures using LISREL 8.54 (Jöreskog & Sörbom, 1993) with maximum likelihood estimation in an exploratory fashion. The competition and practice subscales were analysed separately using the same sequential approach as in the pilot study. This approach first tests separate single factor models for each subscale in order to assess the convergent validity of the items making up that subscale. The goodness of fit of each pair of subscales is then examined to identify any ambiguity among items. Finally, the whole model is tested. At each stage in the process, items may be removed if the fit statistics are inadequate, closer examination of the item reveals some previously undetected theoretical ambiguity, and the item *fails* to meet one or more of the following criteria: 1) the item has a substantial factor loading (greater than 0.40 might be deemed a reasonable minimum) on the hypothesised factor; 2) the standardised residuals associated with the item are modest (in the authors' experience, standardised residuals less than 3.00 might be considered reasonable); and 3) the modification indices associated with the item are not too high (it is difficult to specify precise cut-offs here, but in the authors' experience modification indices in double figures are usually worth exploring further). Of course, factor structures obtained by this exploratory method always require confirmation on a separate sample. Further details of this approach, together with an example of its use, can be found in Rees, Hardy, and Ingledew (2000).

The goodness of fit of models was assessed using the following combination of fit statistics: the Satorra–Bentler chi-square statistic (Satorra & Bentler, 1994) the standardised root mean square residual (SRMR); the root mean square error of approximation (RMSEA; Steiger, 1990); the non-normed fit index (NNFI; Tucker & Lewis, 1973), and the comparative fit index (CFI; Bentler, 1990). This choice of fit statistics was informed by Jaccard and Wan's (1996) recommendation to utilise fit indices taken from three different classes (absolute fit, absolute fit with penalty clauses, and incremental or comparative fit), together with Hu and Bentler's (1999) conclusions regarding the inadequacy of the Goodness of Fit Index. The χ^2 statistic was used as a subjective index of fit (Jöreskog & Sörbom, 1993), with large chi-square values relative to degrees of freedom indicating a poor fit, and small values indicating a good fit (Jöreskog & Sörbom, 1989). However, χ^2 is quite sensitive to sample size and, in particular, gives inflated values for large samples. Consequently, in line with Hu and Bentler's (1999) recommendations, the primary indices of fit were the SRMR, the RMSEA, the NNFI, and the CFI, with cut-off values close to 0.08 for SRMR, 0.06 for RMSEA, and 0.95 for NNFI and CFI being regarded as indicating a good fit for any given model. These are quite stringent criteria.

A slight modification from the pilot study in terms of analytical strategy involved using structural equation modelling in an exploratory fashion; that is to say, at the first and second stages of analysis, items were removed from the extended subscales using standardised residuals, modification indices, and theoretical considerations whenever the fit indices did not meet the cut-off criteria. This process was discontinued if subscales only had four items in them. It resulted in reduced item pools of 44 competition items and 37 practice items for the full model analyses. These item pools were further reduced to four items per subscale following initial tests of the full models, again using a combination of fit statistics, standardised residuals, modification indices, and theoretical considerations. Finally, Cronbach's alphas were computed for each subscale. The main reasons for reducing subscales to four items were: 1) this was the number of items per subscale in the original TOPS; 2) Cronbach's alpha is notoriously problematic with small numbers of items so that three items would likely have caused problems; 3) the larger the number of items per subscale, the more time the instrument would take to complete and the more likely it would be that TOPS data would be contaminated by motivation effects.

Results

Competition subscales

The single factor models for goal setting, imagery, self-talk, and negative thinking only had four items in them. The models for goal setting, self-talk, and negative thinking met or exceeded the fit criteria with the exceptions that χ^2 was high for both self-talk and negative thinking (values of 9.15 and 10.65 with 2 *df*s, respectively). The single factor model for imagery had a very high $\chi^2(2) = 18.31$ and a high RMSEA = 0.09, but otherwise also had an acceptable fit. The item reduction procedure described above was used to reduce the number of items in relaxation, emotional control, distractibility, and automaticity from six to four. The single factor models for these subscales then met or exceeded the fit criteria. The single factor model for activation met the fit criteria with all six items retained.

During the tests of paired models, the activation subscale was reduced to four items. Of the 36 possible pairwise models, 27 met or exceeded the specified criteria. Of the remaining nine models: six models had rather high χ^2 values (ranging from 57.27 to 64.87 with 19 *df*s), but exceeded all of Hu and Bentler's (1999) cut-off criteria; one model had a high χ^2 (19) = 79.90, together with a rather low NNFI = 0.92; one model had a high χ^2 (19) = 111.35 and a high RMSEA = 0.10; and one model had an inadequate fit (χ^2 (19) = 152.20, SRMR = 0.07; RMSEA = 0.11; NNFI = 0.92; and CFI = 0.95). The problem pairings were all associated with self-talk and negative thinking.

The initial fit for the nine-factor model containing all 44 competition items was adequate (χ^2 (866) = 2076.98; SRMR = 0.07; RMSEA = 0.05; NNFI = 0.97; and CFI = 0.97). After item reduction, the fit for the nine-factor model containing four items per subscale was good (χ^2 (558) = 1089.62; SRMR = 0.05; RMSEA = 0.04; NNFI = 0.98; and CFI = 0.98). In light of the conceptual overlap that exists between self-talk and negative thinking, together with the fit problems experienced with the paired models including these subscales, two further eight-factor models were tested. In the first of these models, negative thinking was removed; the fit was very good (χ^2 (436) = 797.83; SRMR = 0.05; RMSEA = 0.04; NNFI = 0.95; and CFI = 0.95). In the second model, self-talk was removed; the fit also remained very good, but not quite as good as when negative thinking was removed (χ^2 (436) = 806.73; SRMR = 0.05; RMSEA = 0.04; NNFI = 0.94; and CFI = 0.95).

Table 1 shows the factor loadings and error terms for the four items comprising each competition subscale. The factor loadings were generally very good, but three items had low factor loadings – one on the automaticity subscale, and two on the distractibility subscale. Means and standard deviations for the competition subscales are provided in Table 3, and the correlations between subscales are shown in Table 4. At this point, the reader should note that the items retained in the emotional control subscales are all negatively phrased and so have negative factor loadings.

Table 1
Completely standardised item loadings and error terms for competition subscales in Study 1.

Competition subscale	Factor loading	Theta-Delta error terms
Factor 1: Self-talk		
Talk positively to get the most out of competitions	0.80	0.37
Manage self-talk effectively	0.76	0.42
Say things to help competitive performance	0.73	0.46
Say specific cue words or phrases to help performance	0.57	0.67
Factor 2: Emotional control		
Emotions get out of control under pressure	−0.84	0.29
Difficulty with emotions at competitions	−0.84	0.29
Difficulty controlling emotions if I make a mistake	−0.80	0.36
Emotions keep me from performing my best	−0.76	0.43
Factor 3: Automaticity		
Able to trust my body to perform skills	0.73	0.46
Sufficiently prepared to perform on automatic pilot	0.61	0.63
Allow whole skill or movement to happen naturally without concentrating on each part	0.47	0.78
Unable to perform skills without consciously thinking	−0.24	0.94
Factor 4: Goal setting		
Set personal performance goals	0.83	0.31
Set very specific goals	0.81	0.34
Evaluate whether I achieve competition goals	0.67	0.55
Set specific result goals	0.59	0.65
Factor 5: Imagery		
Rehearse performance in my mind	0.86	0.26
Imagine competitive routine before I do it	0.83	0.31
Rehearse the feel of performance in my imagination	0.75	0.44
Visualise competition going exactly the way I want it	0.65	0.58
Factor 6: Activation		
Can get myself "up" if I feel flat	0.74	0.45
Can psych myself to perform well	0.74	0.46
Can get my intensity levels just right	0.73	0.47
Can get myself ready to perform	0.66	0.56
Factor 7: Relaxation		
Use relaxation techniques to improve performance	0.89	0.21
Use relaxation strategies as a coping strategy	0.82	0.32
If I'm starting to "lose it", I use a relaxation technique	0.76	0.42
Relax myself to get ready to perform	0.61	0.63
Factor 8: Negative thinking		
Keep my thoughts positive	−0.77	0.40
Self-talk is negative	0.70	0.50
Thoughts of failure	0.69	0.52
Imagine screwing up	0.65	0.58
Factor 9: Distractibility		
Visual distractions would affect my performance	0.76	0.42
Environmental conditions affect my performance	0.49	0.76
My performance would be impaired by sleep loss	0.32	0.90
Loud noises would not affect my performance	−0.09	0.99

Practice subscales

The single factor models for goal setting, imagery, self-talk, and attentional control only had four items in them. The model for imagery exceeded the fit criteria. The single factor models for goal setting, self-talk, and attentional control all had rather high values for both χ^2 (ranging from 8.84 to 14.00 with 2 *df*s) and RMSEAs (ranging from 0.08 to 0.10), but otherwise also exceeded the fit criteria. The single factor models for relaxation, activation, emotional control, and automaticity all exceeded the fit criteria without removing any items.

During the tests of paired models, the relaxation, activation, emotional control, and automaticity subscales were reduced to four items. Of the 28 possible pairwise models, 16 met or exceeded the specified criteria. Of the remaining 12 models: nine models had rather high χ^2 values (ranging from 56.23 to 74.18 with 19 *df*s), but exceeded all of Hu and Bentler's (1999) cut-off criteria; and the remaining three models had high χ^2 values (ranging from 81.51 to 110.99 with 19 *df*s), together with rather high RMSEAs (ranging from 0.08 to 0.09). There was no pattern to these problematic pairings.

The initial fit for the eight-factor model containing all 37 practice items was adequate (χ^2 (601) = 1605.89; SRMR = 0.08; RMSEA = 0.06; NNFI = 0.95; and CFI = 0.95). After item reduction, the fit for the eight-factor model containing four items per subscale was substantially improved (χ^2 (436) = 1009.98; SRMR = 0.06; RMSEA = 0.05; NNFI = 0.96; and CFI = 0.97), with only χ^2 a little high.

Table 2 shows the factor loadings and error terms for the four items comprising each practice subscale. There were very good

Table 2
Completely standardised item loadings and error terms for practice subscales in Study 1.

Practice subscale	Factor loading	Theta-Delta error terms
Factor 1: Self-talk		
Motivate myself to train through positive self-talk	0.79	0.37
Talk positively to get the most out of practice	0.76	0.42
Manage self-talk effectively	0.69	0.52
Say things to myself to help my practice performance	0.61	0.63
Factor 2: Emotional control		
Trouble controlling emotions when things are not going well	−0.83	0.31
Performance suffers when something upsets me	−0.69	0.53
Emotions keep me from performing my best	−0.66	0.56
Frustrated and emotionally upset when practice does not go well	−0.59	0.65
Factor 3: Automaticity		
Able to perform skills without consciously thinking	0.78	0.39
Perform automatically without having to consciously control each movement	0.74	0.46
Allow whole skill or movement to happen naturally without concentrating on each part	0.65	0.58
Monitor all the details of each move to successfully execute skills	−0.41	0.82
Factor 4: Goal setting		
Very specific goals	0.82	0.33
Set goals to help me use practice time effectively	0.79	0.37
Set realistic but challenging goals	0.73	0.47
Don't set goals for practices, just go out and do it	−0.70	0.51
Factor 5: Imagery		
When I visualise my performance, I imagine what it will feel like	0.71	0.50
When I visualise my performance, I imagine watching myself as if on a video replay	0.64	0.59
Rehearse my performance in my mind	0.59	0.65
Visualise successful past performances	0.55	0.70
Factor 6: Activation		
Can get my intensity levels just right	0.70	0.51
Can get myself "up" if I feel flat	0.63	0.61
Can psych myself to perform well	0.56	0.68
I have difficulty getting into an ideal performance state	−0.50	0.75
Factor 7: Relaxation		
I use relaxation techniques to improve my performance	0.82	0.32
Use practice time to work on relaxation technique	0.76	0.42
Practice using relaxation techniques at workouts	0.74	0.45
I use workouts to practice relaxing	0.68	0.54
Factor 8: Attentional control		
Able to control distracting thoughts when training	0.73	0.47
Focus attention effectively	0.70	0.52
Trouble maintaining concentration during long practices	−0.66	0.56
Attention wanders while training	−0.65	0.57

item loadings on all factors, with none that would be considered problematic. Subscale means and standard deviations are provided in Table 3, and the correlations between subscales are shown in Table 4. The reader should again note that the items retained in the emotional control subscale are all negatively phrased.

Cronbach's alphas

Cronbach's alphas for each subscale of TOPS 2 are presented in Table 3. The criterion that is most commonly used for internal consistency is an alpha of 0.70, although Loewenthal (2001) has suggested that an alpha of 0.60 is acceptable for subscales containing only four items. As can be seen from Table 3, the Cronbach's alpha for distractibility/resistance to distraction in the competition subscales was very low ($\alpha = 0.44$). Consequently, a third eight-factor model was run for the competition subscales with distractibility removed. The fit was adequate (χ^2 (436) = 918.03; SRMR = 0.06; RMSEA = 0.04; NNFI = 0.93; and CFI = 0.94), but not as good as the nine-factor model or the other two eight-factor models when either negative thinking or self-talk was removed. Because of this problem with distractibility, two seven-factor models were run with both distractibility and negative thinking removed, and then with distractibility and self-talk removed. The fit for both of these models was good (χ^2 (329) = 639.54; SRMR = 0.05; RMSEA = 0.04; NNFI = 0.95; and CFI = 0.96; and χ^2 (329) = 649.67; SRMR = 0.06; RMSEA = 0.04; NNFI = 0.95; and CFI = 0.95, respectively). The fit was slightly better when distractibility and negative thinking were removed than when distractibility and self-talk were removed.

Table 3
Subscale means, standard deviations, and Cronbach's alphas from Study 1.

Subscale	Competition			Practice		
	M	SD	α	M	SD	α
Self-talk	3.42	0.86	0.82	3.35	0.80	0.82
Emotional Control	3.69	0.83	0.89	3.42	0.76	0.80
Automaticity	3.52	0.63	0.62	3.47	0.65	0.74
Goal Setting	3.73	0.86	0.83	3.24	0.89	0.84
Imagery	3.49	0.94	0.86	3.02	0.82	0.72
Activation	3.84	0.70	0.83	3.41	0.63	0.71
Relaxation	2.80	0.94	0.87	2.18	0.83	0.85
Negative Thinking	2.23	0.77	0.80			
Distractability	2.71	0.68	0.44			
Attentional Control				3.44	0.66	0.78

Finally, because the sample contained participants from three different cultures, we ran sample invariance tests to examine whether the factor structures remained invariant across each pair of samples. Unfortunately, the size of the North American sample was insufficient to run meaningful tests on that sample, so we were constrained to examining sample invariance across the British and Australian samples. For the competition subscales, the test of factor pattern invariance on the eight-factor model with distractibility removed yielded a very good fit, with χ^2 (872) = 1134.44; SRMR = 0.07; RMSEA = 0.04; NNFI = 0.96; and CFI = 0.97. The test for invariance of factor loadings also yielded a very good fit, but the χ^2 difference test indicated that this fit was significantly poorer than the fit obtained when the factor loadings were not constrained to be equal, χ^2 (24) = 324.13, $p < 0.01$. Nevertheless, it was noteworthy that the highest difference between any pair of factor loadings for the Australian and British samples was 0.07. For the practice subscales, the test of factor pattern invariance on the eight-factor model yielded an acceptable fit, with χ^2 (872) = 1249.98; SRMR = 0.07; RMSEA = 0.04; NNFI = 0.93; and CFI = 0.94. The test for invariance of factor loadings also yielded an acceptable fit, but the χ^2 difference test again indicated that this fit was significantly poorer than the fit obtained when the factor loadings were not constrained to be equal, χ^2 (24) = 47.27, $p < 0.01$. Nevertheless, the highest difference between any pair of factor loadings for the Australian and British samples was 0.05.

Discussion

The aim of Study 1 was to examine the factor structure of the TOPS 2 inventory. The fit statistics for all versions of the competition inventory were acceptable. The fit for the eight-factor practice inventory was also good. There were some small questions regarding the possibility of measurement error in the imagery in competition subscale, and χ^2 was a little high in a number of the models tested. Having said that, if only the four fit statistics recommended by Hu and Bentler (1999) are used, then of 380 fit statistics obtained from 90 models tested in Study 1, only 7 fail to meet the very rigorous criteria recommended by Hu and Bentler.

Table 4
Correlations between scores on the competition and practice subscales in Study 1.

	1	2	3	4	5	6	7	8	9	10
1. Self-talk	**0.74**	0.09	−0.02	0.55	0.47	0.46	0.47			0.38
2. Emotional control	0.26	**0.65**	0.19	0.09	−0.04	0.30	−0.01			0.30
3. Automaticity	0.18	0.34	**0.52**	−0.09	0.01	0.16	−0.02			0.06
4. Goal setting	0.49	0.19	0.17	**0.59**	0.44	0.56	0.42			0.48
5. Imagery	0.56	0.22	0.20	0.60	**0.66**	0.37	0.40			0.30
6. Activation	0.50	0.54	0.41	0.47	0.52	**0.63**	0.30			0.61
7. Relaxation	0.54	0.18	0.17	0.37	0.44	0.37	**0.73**			0.23
8. Negative thinking	−0.37	−0.68	−0.33	−0.25	−0.27	−0.53	−0.28			
9. Distractibility	−0.17	−0.36	−0.32	−0.26	−0.33	−0.41	−0.22	0.33		
10. Attentional control	–	–	–	–	–	–	–	–	–	

Note. Correlations among competition subscales are in the lower left diagonal, and those for practice subscales are in the upper right diagonal. Correlations greater than 0.08 are statistically significant ($p < 0.05$, two-tailed).

Cronbach's alpha reliability analysis identified low alphas for two subscales, distractibility in competition ($\alpha = 0.44$) and automaticity in competition ($\alpha = 0.62$). Consideration of the factor loadings given in Table 1 provides an explanation of these results, both subscales contain an item that has a very low loading on the factor. However, the distractibility in competition subscale contained a second item with a low factor loading. Furthermore, this subscale was originally conceived in terms of athletes being resistant to, or distracted by, one or more of a broad range of situational circumstances. The item reduction process resulted in the deletion of many of these different situations, which may have resulted in a subscale that did not adequately sample the population of all possible distractions and was therefore insufficiently focused on the construct. Although the nine-factor model had an acceptable model fit, the fits of the models with distractibility removed were also acceptable, and thus it was felt that a reduced length questionnaire might be more appealing to athletes, given that they are known to dislike lengthy paperwork (cf. Beckmann & Kellmann, 2003). The automaticity subscale was retained for further examination in a second study that was performed to confirm the factor structure obtained in Study 1.

Study 2

Method

Participants

Participants were 277 Australian male (M age = 21.51, SD = 4.94 years) and female (M age = 22.88, SD = 5.74 years) athletes of varying ability from 17 different sports. Participants completed the questionnaire away from the competition and training environments in small groups of 5–20 participants. Responses indicated that 36.10% were competing internationally and 34.70% were competing nationally. Of the rest, 12.30% were competing at a regional or junior national standard, and 11.5% were competing at club or school level. A small percentage (5.40%) of participants failed to report their performance level. All participants gave their informed consent to participate in the study.

Measures

The 64-item TOPS 2, with eight competition and eight practice subscales, from Study 1 was administered.

Procedure

As in Study 1, the TOPS 2 was administered to participants either individually or in small groups of 10–20 people. No other measures were administered with the TOPS 2.

Analytical strategy

Confirmatory factor analysis was again employed to test the factorial validity of the TOPS 2. Specifically, eight-factor CFAs were performed on the competition and practice subscales separately. No sequential testing was performed. The fit criteria from Study 1 were used.

Results

Competition subscale

The fit of the eight-factor model was very good (χ^2 (436) = 695.16; SRMR = 0.06; RMSEA = 0.05; NNFI = 0.97; and CFI = 0.97). Factor loadings and error terms were generally very good and were similar to those obtained in Study 1,[1] although one of the automaticity items (I am unable to perform skills without consciously thinking) displayed a factor loading in the opposite direction, albeit of a non-significant magnitude, to the loading obtained for this item in Study 1. Because of this, we tested an alternative model which removed the final item from the automaticity in competition subscale, thereby leaving only three items to measure automaticity. The fit was very good (χ^2 (406) = 654.37; SRMR = 0.07; RMSEA = 0.05; NNFI = 0.97; and CFI = 0.97) and Cronbach's $\alpha = 0.70$ ($\alpha = 0.67$ for the four item subscale). The correlation between the four item measure of automaticity and the three item measure of automaticity was 0.93.

In Study 1, conceptual overlap between negative thinking and self-talk was highlighted, and models were tested with each of these constructs were removed. To provide continuity with the first study, the same approach was adopted here. The first model removed negative thinking and revealed a good model fit (χ^2 (329) = 484.23; SRMR = 0.06; RMSEA = 0.04; NNFI = 0.97; and CFI = 0.97). Because of the difficulties with the automaticity in competition subscale, we re-ran this model with the final item of automaticity removed. The fit was again good fit (χ^2 (303) = 447.03; SRMR = 0.06; RMSEA = 0.04; NNFI = 0.97; and CFI = 0.97).

In the second model, self-talk was removed. Again the fit of the model was good (χ^2 (329) = 520.27; SRMR = 0.06; RMSEA = 0.05; NNFI = 0.97; and CFI = 0.97), but, as in Study 1, not quite as good as when negative thinking was removed and again we re-ran the model with the final item of automaticity removed. The fit

[1] To avoid unnecessary repetition we have not included the factor loadings and error terms but can do so on request. We also calculated means, standard deviations and Cronbach's alphas for each competition subscale, as well as assessing the correlations between subscales. The results obtained were comparable to that found in Study 1 (e.g., all the Cronbach's alphas were above 0.70 with the exception of automaticity which had an alpha of 0.67). Given the similarity between the results of this study and those of Study 1 we have not included these results to maintain succinctness of the manuscript.

remained good (χ^2 (303) = 483.86; SRMR = 0.06; RMSEA = 0.05; NNFI = 0.97; and CFI = 0.97), but not quite as good as the fit for negative thinking removed.

Practice subscale

The fit for the eight-factor practice model was very good (χ^2(436) = 603.39; SRMR = 0.06; RMSEA = 0.04; NNFI = 0.96; and CFI = 0.96). Factor loadings and error terms were computed, and were comparable with those obtained in Study 1.[2]

Discussion

The initial aim of the present research was to examine the structural integrity of the TOPS using confirmatory factor analysis. The results of the pilot study confirmed some of the problems identified by Lane et al. (2004), notably problems with activation, emotional control, and negative thinking for the competition inventory, and with activation and automaticity for the practice inventory. Furthermore, the present authors had already identified some conceptual problems with the measurement of emotional control in competition and relaxation in practice. All of these problems have been eradicated in the TOPS 2.[3] Furthermore, the factor structure of TOPS 2 has been shown to be invariant across samples drawn from Great Britain and Australia. Nevertheless, the present series of studies identified a new problem with the automaticity in competition subscale. This problem has not been eradicated and is discussed in more detail in the next paragraph. Furthermore, an attempt to introduce an additional subscale measuring resistance to disruption/distractibility in competition failed.

Although the fit statistics reported across studies 1 and 2 are generally quite good, there remain a number of limitations with the TOPS 2. First, the automaticity subscale clearly still requires some attention. The Cronbach's alphas and factor loadings indicate a need to re-examine and probably replace the last item on this subscale. This suggestion seems further vindicated by the difference in the direction of the factor loadings on this item between Study 1 and Study 2 (although the factor loadings in both cases were "non-significant"). The most obvious explanation for this difference in direction is that it is due to what amounts to a double-negative in the item wording, which may lead to confusion from respondents as to how to interpret the item. Re-phrasing the item to remove this double-negative may help to improve both the factor loading and scale reliability. Second, there exists the distinct possibility of conceptual overlap between the subscales of self-talk, negative thinking, and distractibility. Any such conceptual overlap needs to be examined and resolved in future versions of the TOPS. Finally, the TOPS 2, like the TOPS before it, remains a fairly "blunt" instrument. The original authors' intention was to provide an integrated test of psychological skills that might be used in conjunction with other methods to provide an overview of athletes' usage of a broad range of psychological skills. As such, the TOPS may be very appropriate for a lot of applied purposes where it is likely to be supplemented by other forms of assessment. However, it may be completely inappropriate for examining a number of research or applied questions. For example, the TOPS does not differentiate between athletes' use of outcome, performance and process goals (cf. Kingston & Hardy, 1997). Neither does it distinguish between the use of internal visual imagery, external visual imagery, and kinaesthetic imagery (cf. Roberts, Callow, Hardy, Markland, & Bringer, 2008). In the authors' opinion, this is always likely to be a limitation of the TOPS. It is difficult to see how one could succinctly assess such a wide range of psychological skills in an in-depth fashion.

In terms of future research, there is clearly a need to confirm the present findings and further develop the TOPS 2, particularly with regard to the automaticity subscale in the competition inventory. Furthermore, if the distractibility subscale is to be reinstated in future versions of the TOPS then this subscale will require considerable attention. Alternatively, it may be that some sort of attentional control subscale can be resurrected for competition. However, it is not at all clear to the present authors how this might be best achieved (see the earlier discussion on Instrument refinement). It is our view that attentional control is best reflected by its absence in competition, which appears to result in negative thinking.

In light of the fact that the TOPS 2 has already proved popular in the applied domain, there is also a need for some sort of test manual. It is the authors' intention to make such a manual available on-line. Continuing the applied theme, several researchers (e.g., Frey et al., 2003; Thomas et al., 1999) have suggested that practising psychological skills during training might have highly beneficial effects for athletes in competition. This suggestion is worthy of further research.

Perhaps more interesting is the possibility that different psychological skills may be differentially important for different personalities. Personality research seems to have been out of fashion in sport psychology for a number of years (cf. Hardy et al., 1996; Vealey, 2002), and research that examines the interactive effects of personality with psychological skills is probably long overdue. It is entirely plausible that some psychological skills are more important for some people than others. Such research might examine these interactive effects in the context of both training and competition behaviours. In a similar vein, some psychological skills may be more important in some situations than others, but there is a dearth of research that examines the interactive effects of psychological skills usage with either personality or situational factors. These are both worthy objectives for future research. In a rather different direction, the TOPS has probably reached a stage in its development where it might be appropriate to consider translating it into different languages in order that further research might be conducted in other languages, and cultural differences examined.

The present results have a number of applied implications. The eight-factor model (with distractibility removed) for competition has a good fit, but the seven-factor models also display acceptable fits, so the practitioner should feel free to choose the version of the inventory that he or she feels is most appropriate for his/her needs. In the present authors' experience, feedback from practitioners provides strong support for retention of the negative thinking subscale.

In recent years, several researcher-practitioners (e.g., Frey et al., 2003; Muscat, 2004) have summed the practice subscales and the competition subscales of the TOPS to provide overall measures of athletes' use of psychological skills and strategies in these two environments. The authors would like to emphasise that the CFAs reported in the present paper provide support for the use of separate subscale profiles for practice and competition. The analyses do not provide any support for summing subscales to obtain overall scores for competition and practice. We therefore recommend that practitioners use separate subscale profiles rather than global scores for the two domains.

[2] Factor loadings and error terms for the practice subscale can be obtained by contacting the authors. As with the competition subscale, we calculated means, standard deviations and Cronbach's alphas for each practice subscale along with a comparison of the correlations between practice subscales. Results obtained were extremely similar to Study 1.

[3] Copies of the TOPS 2 questionnaire together with instructions regarding its administration and scoring can be downloaded from www.TOPSfirst.com.

A further word of caution is in order regarding comparisons between the use of psychological skills and strategies in competition and the use of those skills and strategies in training. Although Table 4 shows that, as might be expected, there are moderate to strong correlations across contexts, some caution is required in *directly* comparing the different subscales from the TOPS 2 with regard to levels of psychological skills usage at practice and in competition (cf. Frey et al., 2003; Muscat, 2004). Athletes may use different psychological skills and strategies for different purposes in competition compared with training. Furthermore, the subscales may not be measured on a common metric and comparisons may be confounded by response sensitivity and other item differences.[4]

In summary, with the possible exception of one item, the TOPS 2 has quite strong psychometric properties. It can be used as a research tool to examine a number of interesting research questions in the prediction of important training and competition behaviours. It has also been found to be useful in applied settings (e.g., Gould et al., 2002; Lowther et al., 2002) both for profiling athletes' strengths and weaknesses so that interventions can be appropriately targeted, and for assessing the benefits of those interventions.

Acknowledgements

The authors would like to thank Kate Sheppard for her assistance with the data collection for Study 1.

References

Beckmann, J., & Kellmann, M. (2003). Procedures and principles of sport psychological assessment. *The Sport Psychologist, 17*, 338–350.
Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 103*, 411–423.
Biddle, S. J. H., Markland, D., Gilbourne, D., Chatzisarantis, N. L. D., & Sparkes, A. (2001). Research methods in sport and exercise psychology: quantitative and qualitative issues. *Journal of Sports Sciences, 19*, 777–809.
Bond, J., & Sargent, G. (2004). Concentration skills in sport: an applied perspective. In T. Morris, & J. Summers (Eds.), *Sport psychology: Theory, applications and issues* (2nd ed.). (pp. 388–422) Australia: Brisbane: John Wiley & Sons.
Cattell, R. B. (1949). *Manual for the 16 personality factor questionnaire*. Champaign, IL: The Institute for Personality and Ability Testing.
Durand-Bush, N., Salmela, J. H., & Green-Demers, I. (2001). The Ottawa mental skills assessment tool (OMSAT-3). *The Sport Psychologist, 15*, 1–19.
Fletcher, D., & Hanton, S. (2001). The relationship between psychological skills usage and competitive anxiety responses. *Psychology of Sport and Exercise, 2*, 89–101.
Ford, S. K., & Summers, J. J. (1992). The factorial validity of the TAIS attentional-style subscales. *Journal of Sport & Exercise Psychology, 14*, 283–297.
Frey, M., Laguna, P. L., & Ravizza, K. (2003). Collegiate athletes' mental skill use and perceptions of success: an exploration of the practice and competition settings. *Journal of Applied Sport Psychology, 15*, 115–128.
Gauvin, L., & Russell, S. J. (1993). Sport-specific and culturally adapted measures in sport and exercise psychology research: issues and strategies. In R. N. Singer, M. Murphey, & L. K. Tennant (Eds.), *Handbook of research on sport psychology* (pp. 891–900). New York: Macmillan.
Gould, D., Dieffenbach, K., & Moffett, A. (2002). Psychological characteristics and their development in Olympic champions. *Journal of Applied Sport Psychology, 14*, 172–204.
Hardy, L., Jones, G., & Gould, D. (1996). *Understanding psychological preparation for sport: Theory and practice of elite performers*. Chichester: Wiley.
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.
Jaccard, J., & Wan, C. K. (1996). *LISREL approaches to interaction effects in multiple regression*. Thousand Oaks, CA: Sage.
Jackson, S. A., Thomas, P. R., Marsh, H. W., & Smethurst, C. J. (2001). Relationships between flow, self-concept, psychological skills, and performance. *Journal of Applied Sport Psychology, 13*, 129–153.
Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). Newbury Park, CA: Sage.
Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7: user's reference guide*. Mooresville, IN: Scientific.
Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8 user's reference guide*. Chicago, IL: Scientific Software International.
Kingston, K. M., & Hardy, L. (1997). Effects of different types of goals on processes that support performance. *The Sport Psychologist, 11*, 277–293.
Lane, A. M., Harwood, C., Terry, P. C., & Karageorghis, C. I. (2004). Confirmatory factor analysis of the Test of Performance Strategies (TOPS) among adolescent athletes. *Journal of Sports Sciences, 22*, 803–812.
LeUnes, A. D., & Nation, J. R. (1989). *Sport psychology: An introduction*. Chicago, IL: Nelson-Hall.
Loewenthal, K. M. (2001). *An introduction to psychological tests and scales* (2nd ed.). London: UCL Press.
Lowther, J., Lane, A., & Lane, H. (2002). Self-efficacy and psychological skills during the Amputee soccer World Cup. *Athletic Insight: The Online Journal of Sport Psychology, 4*(2). http://www.athleticinsight.com/Vol4Iss2/SoccerSelfEfficacy.htm.
Mahoney, M. J., Gabriel, T. J., & Perkins, T. S. (1987). Psychological skills and exceptional athletic performance. *The Sport Psychologist, 1*, 181–199.
Martens, R. (1977). *Sport competition anxiety test*. Champaign, IL: Human Kinetics.
McNair, D., Lorr, M., & Droppleman, L. (1971). *Profile of mood states manual*. San Diego, CA: Educational & Industrial Testing Service.
Muscat, A.C. (October 2004). Mental skills training: to use or not to use? In: Poster session presented at the annual conference of the association for the advancement of applied sport psychology, Minneapolis, MN.
Nideffer, R. M. (1976). Test of attentional and Interpersonal style. *Journal of Personality and Social Psychology, 34*, 394–404.
Pribram, K. H., & McGuiness, D. (1975). Arousal, activation and effort in the control of attention. *Psychological Review, 82*, 116–149.
Rees, T., Hardy, L., & Ingledew, D. K. (2000). Performance assessment in sport: justification and confirmatory factor analysis of a measurement instrument for tennis performance. *Journal of Applied Sport Psychology, 12*, 203–218.
Roberts, R., Callow, N., Hardy, L., Markland, D., & Bringer, J. (2008). Movement imagery ability: development and assessment of a revised version of the Vividness of movement imagery questionnaire. *Journal of Sport & Exercise Psychology, 30*, 200–221.
Satorra, A., & Bentler, P. M. (1994). Corrections to test statistic and standard errors in covariance structure analysis. In A. Van Eye, & C. C. Clogg (Eds.), *Analysis of latent variables in developmental research* (pp. 399–419). Newbury Park, CA: Sage.
Schutz, R. W., & Gessaroli, M. E. (1993). Use, misuse, and disuse of psychometrics in sport psychology research. In R. N. Singer, M. Murphy, & L. K. Tennant (Eds.), *Handbook of research on sport psychology* (pp. 901–917). New York: Macmillan.
Smith, R. E., Schutz, R. W., Smoll, F. L., & 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 & Exercise Psychology, 17*, 379–398.
Spielberger, C., Gorsuch, R., & Lushene, R. (1970). *Manual for the State-Trait anxiety inventory*. Palo Alto, CA: Consulting Psychologists Press.
Steiger, J. H. (1990). Structural model evaluation: an interval estimation approach. *Multivariate Behavioral Research, 25*, 173–180.
Thomas, P. R., Murphy, S. M., & 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.
Tkachuk, G., Leslie-Toogood, A., & Martin, G. L. (2003). Behavioral assessment in sport psychology. *The Sport Psychologist, 17*, 104–117.
Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1–10.
Tutko, T. A., Lyon, L. P., & Ogilvie, B. C. (1969). *Athletic motivation inventory*. San Jose, CA: Institute for the Study of Athletic Motivation.
Vealey, R. S. (1988). Future directions in psychological skills training. *The Sport Psychologist, 2*, 318–336.
Vealey, R. S. (2002). Personality and sport behavior. In T. S. Horn (Ed.), *Advances in sport psychology* (2nd ed.). (pp. 43–82) Champaign IL: Human Kinetics.
Vealey, R. S., & Garner-Holman, M. (1998). Applied sport psychology: measurement issues. In J. L. Duda (Ed.), *Advances in sport and exercise psychology measurement* (pp. 433–446). Champaign, IL: Human Kinetics.
Weinberg, R. S., & Gould, D. (2003). *Foundations of sport & exercise psychology* (3rd ed.). Champaign, IL: Human Kinetics.
Williams, J. M., & Krane, V. (2001). Psychological characteristics of peak performance. In J. M. Williams (Ed.), *Applied sport psychology: personal growth to peak performance* (4th ed.). (pp. 137–147) Mountain View, CA: Mayfield.

[4] In fact, the common variance shared between the use of psychological skills across the two domains ranges from 27% to 55%, implying that at least 45% of the variance explained remains specific to one or other domain with even the most strongly correlated skill (self-talk in the present instance).