The Sport Psychologist, 2009, 23, 233-251

Automatic Self-Talk Questionnaire for Sports (ASTQS): Development and Preliminary Validation of a Measure Identifying the Structure of Athletes' Self-Talk

Nikos Zourbanos, Antonis Hatzigeorgiadis, Stiliani Chroni, Yannis Theodorakis, and Athanasios Papaioannou
University of Thessaly

The aim of the present investigation was to develop an instrument assessing the content and the structure of athletes' self-talk. The study was conducted in three stages. In the first stage, a large pool of items was generated and content analysis was used to organize the items into categories. Furthermore, item-content relevance analysis was conducted to help identifying the most appropriate items. In Stage 2, the factor structure of the instrument was examined by a series of exploratory factor analyses (Sample A: $N = 507$), whereas in Stage 3 the results of the exploratory factor analysis were retested through confirmatory factor analyses (Sample B: $N = 766$) and at the same time concurrent validity were assessed. The analyses revealed eight factors, four positive (psych up, confidence, anxiety control and instruction), three negative (worry, disengagement and somatic fatigue) and one neutral (irrelevant thoughts). The findings of the study provide evidence regarding the multidimensionality of self-talk, suggesting that ASTQS seems a psychometrically sound instrument that could help us developing cognitive-behavioral theories and interventions to examine and modify athletes' self-talk.

Stoic philosophers argued that thoughts play a critical role in the formulation of behavior and emotions (Reardon, 1993). The inquiry on the role of thoughts in humans is evident in the works of several philosophers, theorists, and researchers. Early on, Plato in his discourse titled *"Theaetetus or about science"* defined thoughts as "the conversation, which the soul holds with itself" (Plato, trans. 1993). Although the concept of thoughts exists from antiquity, in the contemporary literature it is often conveyed by different terms that vary from theorist to theorist (Guerrero, 2005). More recently, Glass and Arnkoff (1997) described thoughts as "cognitive products . . . often referred to as self-statements, self-talk, automatic thoughts or internal dialogue" (p. 911). In the contemporary sport psychology

Zourbanos, Hatzigeorgiadis, Chroni, Theodorakis, and Papaioannou are with the Dept. of PE and Sport Science, University of Thessaly Karies, 42100 Trikala, Greece.

literature the term that has prevailed is self-talk (ST). Hardy, Hall, and Hardy (2005) described ST as a "multidimensional phenomenon concerned with athletes' verbalizations that are addressed to themselves" (p. 905).

Concerning the methodological approach for investigating one's ST, Vygotsky (1986) wrote that "the area of inner speech is one of the most difficult to investigate" (p. 226). The investigations of ST as a covert phenomenon have encountered ample methodological difficulties and challenges. In the cognitive assessment literature, various methodological approaches have been applied to identify the individuals' cognitive processes or structures (see Amsel & Fichten, 1998; Cacioppo, von Hippel, & Ernst, 1997). These techniques range from concurrent to retrospective evaluations and from unstructured to fully structured procedures (e.g., think-aloud, free association, recording of private speech, random sampling, self-monitoring, videotape thought reconstruction, self-statement inventories, clinical interview and thought listing) with several strengths and weaknesses (Blankstein & Segal, 2003). Dobson and Dozois (2003, p. 5) reported, that "access to cognitions is not perfect, and that people may report cognitive activities on the basis of their likelihood of occurrence rather than their actual occurrence." Moreover, verbal reports may involve cognitive processes that are beyond metaconscious control and thus cannot be described by the individuals. Finally, self-reported cognitions are relied on one's memory, and sometimes the information may have been forgotten or recalled inaccurately (Nisbett & Wilson, 1977). Nevertheless, cognitive processes cannot be assessed through external observation and thus self-reports provide us with 'metacognitive knowledge' which can be used to help us understand individuals' perceptions, motives, and cognitions (Guerrero, 2005).

The thought listing technique has been the most widely used method in clinical psychology for treating various subtypes of phobias, anxiety, and depressed mood (for a review see Cacioppo, von Hippel, & Ernst, 1997). In clinical and counseling psychology, the assessment of automatic thoughts is theoretically associated primarily with anxiety and depression and secondarily with positive automatic cognition. Glass and Arnkoff (1997) in their review of cognitive assessment reported numerous self-statement inventories, such as the Automatic Thoughts Questionnaire (ATQ, Hollon & Kendall, 1980) and the Positive Automatic Thoughts Questionnaire (ATQ-P, Ingram & Wisnicki, 1988). Finally, in the educational psychology area, the Self-Talk Inventory (STI, Burnett, 1996) and the Thought Occurrence Questionnaire (Sarason, Sarason, Keefe, Hayes, & Shearin, 1986) were developed to assess students' positive and negative ST, and the occurrence of interfering thoughts during task performance respectively.

Self-talk Instruments in Sport

In the sport setting, ST was initially investigated through the Psychological Skills Inventory for Sport (PSIS; Mahoney, Gabriel, & Perkins, 1987) which was designed to measure exceptional athletic performance. Thomas, Murphy, and Hardy (1999) assessed ST through the Test of Performance Strategies-2 (TOPS-2), which includes two relevant subscales assessing the use of positive ST as a cognitive strategy and the frequency of negative thoughts during training and competition. However, the gap in the assessment of ST led Hardy et al. (2005) to develop

the Self-Talk Use Questionnaire (STUQ). Based on existing qualitative data (Hardy, Gammage, & Hall, 2001), the instrument was developed to assess *Where, When, What,* and *Why* ST is used. Zervas, Stavrou, and Psychountaki (2007), based on Hardy et al.'s (2001) conceptualization, developed the Self-Talk Questionnaire (S-TQ), an instrument assessing the instructional and motivational functions of ST strategies. Finally, Theodorakis, Hatzigeorgiadis, and Chroni (2008) developed the Functions of Self-Talk Questionnaire (FSTQ), an instrument aiming to identify the likely mechanisms through which ST facilitates performance. Nevertheless, none of the above instruments is actually assessing the content of athletes' ST and its underlying structure. In developing cognitive-behavioral theories and interventions that will aptly modify athletes' ST, one should take into consideration the so often highlighted content of ST. In this line of thinking, a fundamental question is raised that ought to be answered beforehand: *What constitutes the athletes' ST*? Despite the increased attention in ST research (for a review see Hardy, 2006), currently only one validated measure exists that explores the underlying structure of athletes' ST, the Thought Occurrence Questionnaire for Sport (TOQS; Hatzigeorgiadis & Biddle, 2000). The TOQS was developed on the basis of the Thought Occurrence Questionnaire (Sarason et al., 1986) and identifies interfering thoughts athletes experience while performing. The instrument consists of three subscales assessing worries related to performance, thoughts of escape and task-irrelevant thoughts. Even though the TOQS has received considerable psychometric support (Hatzigeorgiadis & Biddle, 2000; Lane, Harwood, & Nevill, 2005), it focuses on negative and irrelevant thoughts and has no references to positive thoughts athletes experience.

The development of an instrument assessing the content and the structure of ST may offer an opportunity to better understand the ST phenomenon and the role of ST in sport performance, and help practitioners identifying and modifying irrational or maladaptive ST. Therefore, the aim of this investigation was to develop and test the validity of a questionnaire suitable for detecting and measuring the underlying structure of athletes' ST, the Automatic Self-Talk Questionnaire for Sport (ASTQS).

The investigation was completed in three stages: In Stage 1 item collection was carried out through the thought listing technique and content analysis was used to organize the large pool of items into categories, while the appropriateness of the items was assessed through an item-content relevance analytic method. In Stage 2, the instrument's factor structure was examined through exploratory factor analysis, whereas in Stage 3 the results of the exploratory factor analysis were retested through CFA and concurrent validity was assessed. Ethical approval was granted from the university's research ethics committee.

Stage 1

The aim of the first stage was to draw together a large pool of items both by developing new items and by adapting already existing items from the TOQS (Hatzigeorgiadis & Biddle, 2000), and to identify the underlying structure. This took place in two parts. Firstly, content analysis was conducted as described by Boyatzis (1998) and secondly, item-content relevance was examined based on Dunn, Bouffard and Rogers' (1999) suggestions.

Method

Participants and Procedure. Participants were 355 athletes (194 males and 161 females) with a mean age of 24.34 (± 5.22) years. They competed in eight different team (basketball, soccer, and volleyball) and individual (swimming, athletics, rowing, weightlifting, and wrestling) sports. At the time of data collection, athletes were active at regional or higher level of competition. In particular, 27.8% were competing at international level, 51.7% athletes were competing at national level, and 21.5% were competing at regional or county level. The mean competitive experience of the participants was 8.79 (± 4.60) years. All participants signed a consent form, while for every athlete under the age of 16 parental consent was obtained. Participants listed after the completion of a competition internal talks and thoughts experienced while performing. The procedure was based on Cacioppo and Petty's (1981) and Glass and Arnkoff's (1997) recommendations on thought listing and retrospective recalls, regarding the instructions provided to athletes asking them to write down their thoughts, as well as the amount of time allowed for participants to report their thoughts (i.e., 3 min). Cacioppo and Petty (1981) recommend a 3-min period for thought listing. They suggested that if the most salient thoughts are desired, which was our intention in this phase of the study, then a brief interval is better than a long one. If the interval is too long participants have the time to generate, select and maybe delete portions of cognitive responses.

Content Analysis. Three sport psychologists with experience in qualitative analyses organized the thoughts into categories independently of each other. For categorizing the self-reported thoughts, content analysis following the process described by Boyatzis (1998) was used. The instruction offered for the classification was to organize the thoughts in any possible system of categories based on the content of the responses. Initially, complex thoughts were broken into single subject thoughts. Key words or phrases conveying a single-subject thought were identified and then gathered together into categories. The emerging categories were to be mutually exclusive, exhaustive, ensure independence, and based on a single classification principle. The three sport psychologists, upon completion of their independent work, convened to discuss the categories until agreement was secured. A priori, it was decided that themes not fitted into a category (i.e., marked as miscellaneous) would be deleted. In addition, any category for which consensus was not reached by all three on the themes to be included and/or on its content description would also be deleted.

Before classifying the themes into categories, judges were asked to screen the pool and eliminate or in some cases edit statements that they perceived as redundant, synonymous, and/or incomprehensible. Furthermore, self-reported statements that referred to swears and single word statements (e.g., yes or no) were eliminated. Finally, sport/event-specific technical features and statements that described specific actions/behaviors (e.g., watch your elbow) were edited to become more general if this was possible (e.g., watch your technique), or otherwise eliminated (e.g., straight back).

Results

A total of 648 statements were initially gathered, including the 17 items from the TOQS (Hatzigeorgiadis & Biddle, 2000). More specifically, 544 positive statements and 104 negative statements expressing thoughts that emerged during competition entered the elimination and categorization processes. After completing the single-subject theme identification and the elimination process (as discussed in the previous paragraph) independently, the three judges convened. Consensus was reached for entering in the classification process 180 positive themes and 89 negative themes. Following, these 269 themes were independently listed into categories. Ten categories were originally proposed by the judges but consensus was reached for eight categories containing a total of 93 themes. More specifically, they agreed upon four categories with positive themes, three categories with negative themes and one category with neutral themes.

Forty-five positive thoughts were classified into four categories: The first category included statements related to enhancing physical performance and preparing mentally (9 themes, e.g., let's go), the second category included statements referring to ones' beliefs about him or herself (16 themes, e.g., I believe in my abilities), the third category included statements referring to concentrating and giving instruction (11 themes, e.g., focus on your technique), and the fourth category was characterized by statements to overcome anxiety symptoms (9 themes, e.g., calm down). A fifth category was initially identified by one judge including thoughts referring to issues of technique and tactics. Items from this category were self-instructions regarding technical and tactical aspect of performance (e.g., focus on the ball, footwork, and concentrate on your opponent) and it was finally agreed to include these statement into the concentration—instruction category.

Thirty nine negative thoughts were classified into three categories: The first category included statements that refer to worry and negative self-evaluation over performance (26 themes, e.g., I'm going to lose), the second included statements of withdrawing (6 themes, e.g., I want to stop), and the third category was characterized by statements that described fatigue or somatic exhaustion (7 themes, e.g., I'm tired). Finally there was a neutral category that was characterized by statements irrelevant to the sport setting (9 themes, e.g., I am hungry). These were similar to the categories of Hatzigeorgiadis and Biddle's (2000) study, with the addition of the somatic fatigue / exhaustion category. At this point, we decided to include the neutral themes into the broader negative thoughts dimension based on the structure, but also the high positive correlations, that have been identified for the respective dimension in the TOQS (Hatzigeorgiadis & Biddle, 2000). A fifth category, originally identified by one judge as social comparison, was dropped. Statements in this category included thoughts of negative self-evaluation involving social comparison (e.g., what other will think of my poor performance, others are better than me) and were finally characterized as worry, following Sarason et al.'s (1986) conceptualization who considered social comparison among the major components of worry.

Method

Item Content Relevance Analysis

The 45 positive and the 48 negative (including the neutral) items that were selected were subsequently tested through a structured content analytic method based on Dunn et al.'s (1999) suggestions. Firstly, we listed the items in two forms (one for the positive and one for the negative items) which were distributed to ten judges (5 academics, 3 coaches, and 2 athletes; all holding higher degrees in relevant areas) that were not involved in the writing phase of the instrument construction process, in order not to be biased in their assessment about item-domain matches. The judges first familiarized themselves with the definitions of the categories provided and then matched each of the 45 positive and 48 negative items with the specified categories on a 5-point rating scale (1 = *poor match*, 2 = *fair match*, 3 = *good match,* 4 = *very good match,* 5 = *excellent match*). More specifically, we asked the panel to classify the positive items into four given categories of psych up, confidence, concentration—instruction, and anxiety control, and the negative items into four categories of worry, disengagement, somatic fatigue, and irrelevant thoughts. We selected the categories in accordance to the content analysis from the thought listing technique regarding the structure of the athletes' ST. Furthermore, we added in each case a fifth category labeled 'other-specify', and we asked judges to indicate whether the item could not be adequately described by any of the listed categories, and subsequently to suggest any category that could be added. Finally, we instructed the judges to include the items into more than one category if they thought this was appropriate and to provide any comments regarding the wording, grammar and content of the items. The purpose of the matching task protocol was to quantify item ratings and to evaluate and summarize judgments using quantitative statistical procedures.

The evaluation of the judges' ratings was based on the content validation procedures (indices) of Aiken's (1985) item-content validity coefficient (V) and Cohen's (1977) effect size (ES) index for dependent means. The V statistic provides the statistical significance of judges' ratings for the key-construct that each item is supposed to measure. The V values can range from 0 to 1 (1 indicating perfect agreement). The V values are then compared against a right-tailed binomial probability table provided by Aiken (for a more detailed discussion see, Aiken, 1985). In addition, and following Dunn et al.'s (1999) recommendations, because the V statistic provides information regarding an item's content-match with one construct (the one that is supposed to measure) and not with all constructs, it was deemed appropriate to examine whether the items were not matching with the other constructs under examination. For that reason, planned contrasts for the positive (n = 4) and the negative (n = 4) constructs between the key domain-construct and the remaining non key-constructs were tested for the 45 positive and 48 negative items. The Cohen's (1977) ES index for dependent means (d_z') computes the size of the difference between two means (the key-construct and each one of the nonkey constructs) for each item. Cohen's (1977) guidelines for ES suggest that values from .50 to .79 are considered to represent moderate ES, whereas values greater than .80 are considered indicative of large ES.

Results

From the positive categories 12 items were excluded with V values lower than .45 (not significant) and ES lower than .47. From the negative categories 17 items were excluded with V values lower than .45 (not significant) and ES lower than .47. The results from the V coefficient showed that these items were not adequately relevant to the constructs they were supposed to measure, whereas the results from the planned contrasts indicated that they were also relevant to other nonkeyed domains. Regarding the items that were retained for the positive categories, V values ranged from .55 to .95 ($p < .05$) and ES ranged from.57–4.69, whereas for the negative categories V values ranged from .63 to .93 ($p < .05$) and ES ranged from .57 to .93[1]. A total of 32 positive items and 31 negative were retained for the next stage of the investigation.

Stage 2

The purpose of this stage was to explore the factorial structure of the instrument. First, separate exploratory factor analyses were conducted for the positive and negative ST dimensions. The solutions that emerged were subsequently tested together.

Method

Participants and Procedure. The instrument was distributed to 507 athletes (284 males and 221 females; mean age 19.87, ± 6.22 years; mean competitive experience 7.58, ± 4.83 years) representing a variety of individual (athletics, swimming, and tennis: $n = 195$) and team (football, basketball, volleyball, handball, and water polo: $n = 310$) sports. With regard to competitive level, 9.3% of them had competed at international level, 50.3% at national level, and 37.5% at regional or county level. Participants were informed that the forms they were going to complete were anonymous and signed a consent form. They were asked to indicate based on their latest competitions thoughts they usually experience or intentionally use while performing. Responses were given on a 5-point scale (0 = *never*, 1 = *rarely*, 2 = *sometimes*, 3 = *often,* 4 = *very often*) based on Amsel and Fichten's (1998) recommendations about end points for self-statements inventories.

Results

Positive ST Dimension. Principal components analysis was initially conducted to identify the number of factors to be retained through the scree plot (Kline, 1994). The analysis revealed that a four- or a five-factor solution should be interpreted. To further clarify the number of factors to be interpreted, factor analysis on both solutions was performed (Floyd & Widaman, 1995). In the first analysis, the five-factor solution was tested using principal axis factoring and oblique rotation, because factors were correlated. The analysis indicated that the four-factor solution should be interpreted because the fifth factor included only one substantive item. Subsequently, a four-factor solution was tested through

principal axis factoring with oblique rotation. Items with loadings lower than .40 and items with similar loadings on more than one factor were eliminated to obtain a clear solution and improve the independence of factors (Kahn, 2006). The final solution included 19 items. The four factors explained 51.94% of the total variance and factor loadings ranged from .43 to .76. Factors were labeled according to the items with the highest loadings. The highest loadings in the first factor were "I feel strong" and "I believe in me" and the factor was labeled 'confidence'. The highest loadings in the second factor were "do your best" and "give 100%" and the factor was labeled 'psych-up'. The highest loadings in the third factor were "focus on your technique" and "concentrate on your goal" and the factor was labeled 'instruction'. Finally, the highest loadings in the fourth factor were "calm down" and "relax" and the factor was labeled 'anxiety control.'

Negative ST Dimension. The same procedures were followed for the negative ST dimension. The scree plot from the initial principal components analysis indicated that a four or a five factor solution should be retained. First, a five factor solution was tested through principal axis factoring and oblique rotation. The analysis revealed that the four-factor solution should be interpreted because the fifth factor included only one substantive item. After eliminating items with loadings lower than .40 and items with similar loadings on more than one factor, the final solution included 21 items. The four factors explained 58.70% of the total variance and factor loadings ranged from .40 to .73. Factors were labeled according to the items with the highest loadings. The highest loadings in the first factor were "I think I'll stop trying" and "I can't keep going", and the factor was labeled 'disengagement'. The highest loadings in the second factor were "I am not going to reach my goal" and "I'm wrong again" and the factor was labeled 'worry'. The highest loadings in the third factor we "I'm hungry" and "I want to take a shower", and the factor was labeled 'irrelevant thoughts'. Finally, the highest loadings in the fourth factor were "my legs/arms are shaking from tiredness" and "my body doesn't help me today", and the factor was labeled 'somatic fatigue.'

Combined Positive and Negative ST. Finally, the solutions that emerged for the two ST dimensions were tested in a single exploratory factor analysis to test whether a 2-general-factor (positive—negative ST) or an 8-factor solution in accordance to the separate analyses would emerge. Principal axis factoring with oblique rotation supported the 8-factor interpretation. The eight factors explained 57.05% of the total variance and factor loadings ranged from .36 to .69. The results of the exploratory factor analysis are presented in Table 1. Descriptive statistics, Pearson's correlations and Cronbach's alpha coefficients between the subscales are presented in Table 2.

Stage 3

The purpose of the final stage was to test on a different sample the factor structure of the instrument that emerged from the previous stage and to provide evidence regarding the concurrent validity of the instrument.

Table 1 Exploratory Factor Analysis for the ASTQS (*N* = 507)

	F1	F2	F3	F4	F5	F6	F7	F8
Let's go	.48							
Power	.49							
Give 100%	.56							
Do your best	.68							
Strong	.49							
Relax		.55						
Don't get upset		.52						
Calm down		.61						
No stress		.43						
I believe in me			.59					
I am very well prepared			.42					
I feel strong			.59					
I can make it			.37					
I believe in my abilities			.55					
Concentrate on your goal				.58				
Focus on what you need to do now				.38				
Concentrate on your game	.32			.48				
Focus on your technique				.69				
Concentrate				.48				
I am going to lose					.52			
I'm wrong again					.60			
I am not as good as the others					.59			
I am not going to reach my goal					.63			
I cannot concentrate					.38			
I am not going to make it					.44			
What will others think of my poor performance					.45			
I want to stop						.60		
I want to get out of here						.59		
I think I'll stop trying						.65		
I can't keep going						.63		
I am fed-up						.56		
My body is not in a good condition							-.50	
I am tired							-.46	
Today I 'suck'							-.36	
My legs/arms are shaking from tiredness							-.60	
My body doesn't help me today							-.54	
I am thirsty								.41
What will I do later tonight								.42
I am hungry								.63
I want to take a shower								.58
Variance %	3.52	3.79	2.94	13.45	4.91	22.66	2.64	3.16

Note. Loadings below .30 have been omitted. F1: psych-up, F2: anxiety control, F3: confidence, F4: instruction F5: worry, F6: disengagement, F7: somatic fatigue, F8: irrelevant thoughts

Method

Participants and Procedure. The instrument was distributed to 766 athletes (466 males and 299 females; mean age 17.44, ± 5.18 years; mean competitive experience 5.22, ± 3.27 years) representing a variety of individual (athletics, swimming, tennis, and rowing*:* $n = 421$) and team (football, basketball, volleyball, and handball: $n = 339$) sports. With regard to competitive level, 10% of them had competed at international level, 41.1% at national level, and 48.4% at regional or county level. Procedures and instructions were similar to those of the second stage.

In addition to the ASTQS, other scales were administered to different parts of the sample to test the concurrent validity of the instrument. In particular, 76 participants completed the Test Of Performance Strategies-2 (TOPS; Thomas et al., 1999) including two general positive and negative ST subscales; 248 participants completed the Sport Anxiety Scale (SAS; Smith, Smoll, & Schutz, 1990) including cognitive anxiety, somatic anxiety and concentration disruption; 112 participants completed the confidence subscale from the Competitive State Anxiety Inventory-2R (CSAI-2R; Cox, Martens, & Russell, 2003); 174 participants completed three subscales from the Affect State Inventory (ASI; Kakkos & Zervas, 1997), vigor, tension, and boredom.

Factorial Validity. The factor structure of the questionnaire that emerged from the exploratory factor analysis was tested through confirmatory factor analysis (CFA). Three fit indices were used to assess the adequacy of the tested models, the Comparative Fit Index (CFI), the Non-Normed Fit Index (NNFI), and the Root Mean Square Error of Approximation (RMSEA). Three alternative models were tested: (a) an 8-factor model where factors were allowed to correlate according to the results of the exploratory factor analysis, (b) a 2-factor model where all positive ST items were set to load on a single positive factor and all negative ST items were set to load on a single negative factor (the two factors were allowed to correlate), and (c) a 10-factor model where the four positive ST factors were set to form a second-order positive ST factor, and the four negative ST factors were set to form a second-order negative ST factor.

Finally, the adequacy of the model across gender was tested. Multisample analyses were performed to test for invariance of factor loadings and correlations across levels of gender. The invariance routine recommended by Byrne, Shavelson and Muthen (1989) was adopted. Initially, a baseline-unconstrained model was estimated to test whether the factor pattern (number of factors and number of indicators) was similar across groups. This was followed by a model where invariance of factor loadings and invariance of factor correlations were tested. The models were evaluated using the same fit indices, as in the single sample analyses. Cheung and Rensfold (2002) suggested that evaluation of models through chi-square differences is likely to give significant differences, due to the sensitivity of chi-square to sample size. Subsequently, they recommend the use of the CFI, with differences of .01 or less in the CFI between the baseline and subsequent restricted models, to be supportive for the equivalence of the fixed parameters across the samples. Furthermore, Marsh, Marco and Asçi (2002) supported that changes in the RMSEA should also be used to evaluate differences in restricted models, with very small changes being substantively unimportant.

Table 2 Descriptive Statistics, Pearson's Correlations and Cronbach's Alpha Coefficients for the ASTQS Subscales in Stages 3 and 4

	Descriptive statistics		*Alpha*	Correlations						
Stage 3 (*N* = 507)	***M***	***SD***		**1**	**2**	**3**	**4**	**5**	**6**	**7**
1 Psych Up	3.09	.69	.75	—						
2 Anxiety Control	2.75	.78	.64	.31**	—					
3 Confidence	2.78	.72	.76	.51**	.26**	—				
4 Instruction	2.87	.78	.79	.46**	.34**	.55**	—			
5 Worry	1.44	.80	.85	-.29**	.01	-.30**	-.05	—		
6 Disengagement	1.01	.89	.84	-.34**	-.06	-.22**	-.09*	.63**	—	
7 Somatic Fatigue	1.53	.74	.74	-.15**	-.03	-.13**	.01	.58**	.60**	—
8 Irrelevant Thoughts	1.23	.95	.75	-.22**	.03	-.06	.04	.53**	.59**	.50**
Stage 4 (*N* = 766)	***M***	***SD***	***Alpha***	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1 Psych Up	3.05	.89	.89	—						
2 Anxiety Control	2.39	.94	.94	.54**	—					
3 Confidence	2.83	.92	.92	.74**	.52**	—				
4 Instruction	2.97	.87	.87	.66**	.53**	.64**	—			
5 Worry	1.05	.83	.83	-.27**	-.21**	-.45**	-.24**	—		
6 Disengagement	.65	.79	.79	-.31**	-.25**	-.32**	-.31**	.65**	—	
7 Somatic Fatigue	1.04	.84	.84	-.23**	-.23**	-.31**	-.20**	.71**	.68**	—
8 Irrelevant Thoughts	.97	.89	.89	-.22**	-.10**	-.15**	-.09*	.53**	.52**	.67**

*$p < .05$, ** $p < .01$

Concurrent Validity. Concurrent validity was examined through correlations between the ASTQS and the various scales that administered to parts of the sample. It was expected that the positive ST subscales from the ASTQS would correlate positively with the conceptually positive subscales from the other instruments (e.g., TOPS-positive, CSAI-2R-confidence, ASI-vigor) and negatively with the conceptually negative subscales from other instruments (e.g., TOPS-negative, SAS-cognitive anxiety, ASI-tension). Similarly, the negative ST subscales from the ASTQS would correlate positively with the conceptually negative subscales from the other instruments (e.g., TOPS-negative, SAS-cognitive anxiety, ASI-tension), and negatively with the conceptually positive subscales from the other instruments (e.g., TOPS-positive, CSAI-2R-confidence, ASI-vigor).

Results

Factorial Validity. Confirmatory factor analysis using EQS 6.1 was conducted to test the models. Model parameters were estimated based on the covariance matrix and using the robust method, because examination of the descriptive statistics revealed small deviations from univariate normality for some of the items (kurtosis greater than 2.0). First, an 8-factor correlated model was estimated with uncorrelated residuals for the indicators. The fit indices (Table 3) supported the adequacy of the model. All items had high loadings and relatively low errors, which in addition to the adequacy of the fit indices support the hypothesized factor structure of the instrument. The fit indices for 2-factor model showed that the model did not adequately represent the data. Finally, the fit indices for the 10-factor model showed a slightly worse fit compared with the eight-factor model, which however was acceptable (Table 3).

Invariance of Model Across Gender. To proceed to the tests of invariance across gender, the 8-factor model was first tested separately for males and females. The fit indices showed acceptable fit. Subsequently, the multisample analyses were calculated. The fit indices for the baseline unconstrained model supported the invariance of factor pattern across gender. The analysis for the constrained model showed that the difference in the CFI between the baseline and the restricted model was less than .01 (.006), whereas no differences in RMSEA were detected, thus providing support for the invariance of factor loadings and factor correlations between males and females. The fit indices for the tests of invariance are presented in Table 3. Descriptive statistics, Pearson's correlations and Cronbach's alphas for all subscales are presented in Table 2.

Concurrent Validity. Concurrent validity was examined through correlations between the ASTQS and the SAS, the CSAI-2R confidence subscale, the ASI and the TOPS (Table 4). Regarding the TOPS, the positive ST subscale correlated positively with the positive ST subscales from the ASTQS. Respectively, the TOPS negative ST subscale correlated positively with the negative subscales from the ASTQS. Regarding the ASI, vigor correlated positively with the positive ST subscales and negatively, with the negative ST subscales from the ASTQS, whereas the tension and boredom subscales correlated negatively the positive ST and positively with the negative ST subscales from ASTQS. The anxiety subscales from the SAS correlated positively with the negative ST subscales. Finally, the

Table 3 Confirmatory Factor Analysis: Fit Indices for the Single-Sample and Multisample Models

Samples	Model	x^2	*df*	CFI	NNFI	RMSEA
Full sample (N = 766)	2-factor model	2459.42**	739	.84	.83	.06
	8-factor model	1404.97**	712	.94	.93	. 04
	10-factor model	1570.93**	731	.92	.92	.04
Males (N = 466)	8-factor model	969.86**	712	.93	.92	.04
Females (N = 299)	8-factor model	1270.20**	712	.93	.92	.05
Multisample	Invariance tests					
	Baseline	2554.12**	1426	.90	89	.03
	Constrained	2679.13**	1485	.89	.89	.03

* $p < .05$, ** $p < .00$

confidence subscale form the CSAI-2R was positively correlated with the positive ST subscales and negatively with the ST subscales from the ASTQS.

Discussion

The aim of the current investigation was to develop a questionnaire assessing athletes' ST. In brief it should be noted that this study is one of the first of its kind to demonstrate the latent structure of athletes' ST. The results can be summarized as follows. First, the responses obtained from the athletes and the classification into categories using a qualitative data approach (Boyatzis, 1998) suggested four positive, three negative and one neutral category. The positive categories were psych up, confidence, instruction, and anxiety control. The negative categories were worry, disengagement and somatic fatigue and the neutral category was irrelevant thoughts. Further, the use of a structured content analytic method based on Dunn et al.'s (1999) suggestions helped in discarding items and identifying those to be considered for the exploratory factor analysis, which yielded an acceptable 8-factor structure. Finally, the psychometric integrity of the instrument was also supported through evidence of construct and concurrent validity. Moreover, reliability analyses provided evidence for the internal consistency of the scale. Altogether, these findings seem to suggest that the ASTQS is a reliable and valid self-report instrument for assessing athletes' trait ST.

Thoughts were categorized as positive or negative based on their wording and not on the results or consequences that can generate. The type of thoughts that were identified has been previously encountered in the sport psychology literature in various studies pertaining to the use and functions of ST. Analytically, in relation to the content analysis four positive ST categories were identified. The first category was labeled psych up and included statements referring to energising (e.g., let's go, go strong) and maximizing effort (e.g., do your best, give 100%), but also more sport-specific cues such as "hit strong," "play strong," and "power legs" which were reworded to improve generalisability. The content of the statements in this category were similar to the statements that were described from

Table 4 Correlations Between the ASTQS, the SAS, CSAI-2R, the ASI, and the TOPS Subscales

	SAS (*N* = 248)			CSAI-2R (*N* = 112)		ASI (*N* = 174)		*TOPS* (*N* = 76)	
ASTQS	**Cognitive Anxiety**	**Somatic Anxiety**	**Concentration Disruption**	**Confidence**	**Vigor**	**Tension**	**Boredom**	**Positive Self-talk**	**Negative Thinking**
Psych Up	-.12	.04	-.08	.32**	.26**	-.16*	-.26**	.56**	-.33**
Confidence	-.34**	.00	-.10	.53**	.44**	-.36**	-.35**	.63**	-.29*
Instruction	-.05	.13*	-.09	.35**	.37**	-.23**	-.24**	.48**	-.13
Anxiety Control	.06	.24**	-.01	.25**	.29**	-.12	-.22**	.24*	-.03
Somatic Fatigue	.38**	.28**	.30**	-.16	-.45**	.51**	.50**	-.12	.16
Disengagement	.40**	.29**	.32**	-.12	-.42**	.39**	.40**	-.25*	.46**
Worry	.50**	.25**	.14*	-.30**	-.54**	.56**	.47**	-.37**	.66**
Irrelevant Thoughts	.14*	.16*	.23**	-.19	-.40**	.48**	.49**	-.03	.14

[1] The full Table comprising statistics for all items was omitted due to its length and complexity; however, it is available from the first author upon request.

*$p < .05$; ** $p < .01$

Weinberg, Gould, and Jackson (1980) who evaluated the positive effects of 'psych-up' techniques in an experimental study to enhance participants' motivation and energy. The second category included statements referring to ones' beliefs about adequacy of abilities (e.g., I believe in me, I'm strong) and state of readiness (e.g., I'm well prepared, I can make it) and was labeled confidence. In the literature there is evidence that athletes use ST to build confidence and instill belief in their abilities for encountering demanding situations (Hardy et al., 2001). The third category was instruction. Regarding the content of the statements, some were general such as "focus" and "concentrate on your game" whereas most were sport-specific instructions such as "concentrate on your backhand", "bend your knees", or "high elbow". These latter statements were edited to become more general such as "concentrate" or "focus on your technique." Nideffer (1993) suggested that ST could help directing and redirecting attention to task relevant cues, and Van Raalte, Brewer, Rivera, and Petipas (1994) reported that athletes' ST frequently involves such cues. Finally, the fourth category was characterized by statements that refer to anxiety control. In this category statements were included instructing athletes achieving desired cognitive and emotional states, helping them not becoming overaroused (e.g., relax, calm down) and overcoming anxiety symptoms (Hanton & Jones, 1999).

Regarding the negative ST three categories were identified. The first category included statements that refer to worry, negative self-evaluation over performance and social comparison, which have been recognized as the most frequent negative thoughts athletes experience (Hatzigeorgiadis & Biddle, 2000). The second category included statements referring to withdrawal symptoms, which is also part of the TOQS. In addition to the previously identified categories, a third category emerged including statements characterized as perceptions of somatic fatigue. Royal, Farrow, Mujika, Halson, Pyne, and Abernethy (2006) acknowledged that athletes usually experience such thoughts during match-play and these can be crucial for performance and decision making. In this category, statements had to do with symptoms of tiredness and unpleasant body symptoms (e.g., I'm tired, my body doesn't help me today). In some instances such statements were also related to thoughts of withdrawal (e.g., my legs are exhausted I can't run anymore, I think I'm going to faint). These items were excluded to increase the independence of the categories. Finally, the neutral category, which was irrelevant thoughts or neutral thoughts, included statements that were completely irrelevant to match-play. This category for the purposes of the analyses was eventually merged with the broader negative ST dimension because evidence has shown large correlations between negative and task-irrelevant thoughts (Hatzigeorgiadis & Biddle, 2000). Overall, through the content analysis and the item-content relevance analysis the most frequent and most representative statements of each dimension were selected, while care was taken to improve generalisability of the items across sports.

Exploratory factor analysis was subsequently calculated to identify the underlying structure of athletes' ST. The analysis yielded eight factors, four positive (psych up, confidence, instruction, anxiety control) and four negative (worry, disengagement, somatic fatigue and irrelevant thoughts). Finally, confirmatory factor analysis provided evidence that the ASTQS is best represented by a model with eight distinct dimensions of thoughts. That both the 8-factor model and the second order 10-factor model were acceptable in describing the structure of the instru-

ment, suggests that the eight factors assess different ST dimensions, which however represents two broader positive and negative dimensions. In contrast, the 2-factor model which included a positive and a negative dimension did not fit the data well. In addition, results regarding the invariance of the model suggest that the questionnaire is equally effective for male and female athletes showing no differences in the structure of thought patterns.

The results obtained in third stage of the study appear to be encouraging regarding the concurrent validity of the ASTQS. Following Kline's (1993) recommendations, correlation analysis was used to test the degree to which scores from the ASTQS subscales related to scales assessing similar constructs, that were applied simultaneously. The pattern of correlations that emerged showed that the positive ST subscales from the ASTQS had positive correlations with positively valenced and negative correlations with negatively valenced scales from other instruments (TOPS, the SAS, the CSAI-2R, and the ASI). Accordingly, the negative ST subscales from the ASTQS had positive correlations with negatively valenced and negative correlations with positively valenced scales from other instruments (TOPS, CSAI-2R, and ASI). Overall, the patterns of correlations that emerged provide adequate support for concurrent validity.

Limitations and Future Validation

One methodological limitation that should be reminded with regard to the present instrument is the use of self-reports. Cognitive processes often occur without conscious awareness, and also rely on memory (Nisbett & Wilson, 1977). Nevertheless, self-reports provide metacognitive knowledge, access to subjective attitudes, experiences and cognitive operations that cannot be obtained through external observation (Guerrero, 2005). In addition, first-hand reports provide potential evidence of ecological validity. Another important issue regarding self-reports is that they are liable to socially desirable responses. The problem of social desirability exists in most assessments based on self-reports. To minimize the likely effects, participants received instructions aiming at reducing socially desirable responses. Nevertheless, further validation could consider the effects of social desirability.

The validation of the ASTQS was based on a trait approach. Given the size of the samples required to complete the appropriate psychometric analyses, it was not possible to collect postcompetition questionnaires. However, the scale could be potentially used as a measure of state ST (thoughts athletes have experienced in a given competition). Therefore, further research could examine the validity of the instrument in assessing situational ST.

With regard to content of the scales, the psychometric theory supports that content validity is related to the number of items representing a construct. The present investigation aimed to identify the more salient items representing the identified constructs, thus relying on data reduction techniques. These techniques compromise the ability of instruments to fully assess a construct, however they are essential in developing scales that can be used effectively. Contemporary measures tend to get smaller to facilitate the independence of factors and to gain brevity and applicability Burisch (1997).

Finally, when examining the psychometric properties of a new instrument it is not feasible to test all types of validity in a single investigation. In the present investigation the concurrent validity of the ASTQS was tested through correlational analyses with scales assessing similar constructs. The concurrent validity support can only be considered preliminary because all the measures share method variance, however it should be noticed that there is no objective criterion to compare against when assessing thoughts. The use of alternative research paradigms, such as the use of the Self-Talk and Gestures Rating Scale (Van Raalte et al., 1994), a measure developed to assess observable ST, could further support aspects of validity. Furthermore, issues of discriminant, predictive and incremental validity could be examined in future research.

Identifying the structure of athletes' ST can help enhancing our understanding regarding the role of thoughts athletes experience during sport competition and advance ST research. Furthermore, it can help athletes, coaches and sport psychologists evaluate with greater precision athletes' thoughts and facilitate developing and applying effective ST plans for athletes. Toward this direction, the ASTQS, an instrument identifying the multidimensional structure of athletes' ST, appears to be a promising tool.

Acknowledgments

This research was supported by the State Scholarships Foundation of Greece awarded to the first author. We would also like to thank all these athletes and coaches, for their help and cooperation.

References

Aiken, L.R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45,* 131–142.

Amsel, R., & Fichten, C.S. (1998). Recommendations for self-statements inventories: Use of valence, end points, frequency, and relative frequency. *Cognitive Therapy and Research, 3,* 255–277.

Blankstein, K.R., & Segal, Z.V. (2003). Cognitive assessment. Issues and methods. In K.S. Dobson (Ed.), *Handbook of cognitive-behavioral therapies* (2nd ed., pp. 40–85). New York: The Guilford Press.

Boyatzis, R.E. (1998). *Transforming qualitative information: Thematic analysis and code development*. London: Sage.

Burisch, M. (1997). Test length and validity revisited. *European Journal of Personality, 11,* 303–315.

Burnett, P.C. (1996). Childrens' self-talk and significant others' positive and negative statements. *Educational Psychology, 16,* 57–68.

Byrne, B.M., Shavelson, R.J., & Muthen, B. (1989). Testing for the equivalence of factor covariance and means structures: The issue of partial measurement invariance. *Psychological Bulletin, 105,* 456–466.

Cacioppo, J.T., & Petty, R.E. (1981). Social psychological procedures for cognitive response assessment: The thought listing technique. In T. Merluzzi, C. Glass, & M. Genest (Eds.), *Cognitive assessment* (pp. 309–342). New York: Guilford.

Cacioppo, J.T., von Hippel, W., & Ernst, J.M. (1997). Mapping cognitive structures and processes through verbal content. The thought-listing technique. *Journal of Consulting and Clinical Psychology, 65*(6)*,* 928–940.

Cheung, G.W., & Rensfold, R.B. (2002). Evaluating goodness of fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233–255.

Cohen, J. (1977). *Statistical power and analysis for the behavioural sciences.* (Rev. ed.) New York: Academic.

Cox, R.H., Martens, M.P., & Russell, W.D. (2003). Measuring anxiety in athletics: The Revised Competitive State Anxiety Inventory–2. *Journal of Sport & Exercise Psychology, 25,* 519–533.

Dobson, K.S., & Dozois, D.J.A. (2003). Historical and philosophical bases of the cognitive-behavioral therapies. In K.S. Dobson (Ed.), *Handbook of cognitive-behavioral therapies* (2nd ed., pp. 4–39). New York: The Guilford Press.

Dunn, J.G.H., Bouffard, M., & Rogers, W.T. (1999). Assessing item content-relevance in sport psychology scale-construction research: Issues and recommendations. *Measurement in Physical Education and Exercise Science, 3,* 15–36.

Floyd, F.J., & Widaman, K.F. (1995). Factor analysis in the development and refinement of clinical assessments instruments. *Psychological Assessment, 7,* 286–299.

Glass, C.R., & Arnkoff, D.B. (1997). Questionnaire methods of cognitive self-statements assessment. *Journal of Consulting and Clinical Psychology, 6,* 911–927.

Guerrero, M.C.M. (2005). *Inner speech-L2: Thinking words in a second language*. New York: Springer.

Hanton, S., & Jones, G. (1999). The effects of a multi-modal intervention program on performers: II. Training the butterflies to fly in formation. *The Sport Psychologist, 13,* 22–41.

Hardy, J. (2006). Speaking clearly: A critical review of the self-talk literature. *Psychology of Sport and Exercise, 7,* 81–97.

Hardy, J., Gammage, K., & Hall, C.R. (2001). A descriptive study of athletes self-talk. *The Sport Psychologist, 15,* 306–318.

Hardy, J., Hall, C.R., & Hardy, L. (2005). Quantifying athlete self-talk. *Journal of Sports Sciences, 23,* 905–917.

Hatzigeorgiadis, A., & Biddle, S.J.H. (2000). Assessing cognitive interference in sports: The development of the Thought Occurrence Questionnaire for Sport (TOQS). *Anxiety, Stress, &. Coping, 13,* 65–86.

Hollon, S.D., & Kendall, P.C. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research, 4,* 383–395.

Ingram, R.E., & Wisnicki, K.S. (1988). Assessment of positive automatic cognition. *Journal of Consulting and Clinical Psychology, 56,* 898–902.

Kahn, J.H. (2006). Factor analysis in counselling psychology, research, training and practice: Principles, advances, and applications. *The Counseling Psychologist, 34,* 684–718.

Kakkos, V., & Zervas, Y. (1997). Validation of the Affect State Inventory (ASI): A preliminary study in competitive situations. In Lidor, R., & Bar-Eli, M. (Eds.) Innovations in sport psychology: Linking theory and practice. Proceedings of IX world congress of sport psychology, Part I, (pp. 339-341). Netanya, Israel.

Kline, P. (1994). *An easy guide to factor analysis*. London: Routledge.

Kline, P. (1993). *The handbook of psychological testing*. London: Routledge.

Lane, A.M., Harwood, C., & Nevill, A.M. (2005). Confirmatory factor analysis of the Thought Occurrence Questionnaire for Sport (TOQS) among adolescent athletes. *Anxiety, Stress, and Coping, 18*(3), 245–254.

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

Marsh, H.W., Marco, I.T., & Asçi, F.H. (2002). Cross-cultural validity of the physical self-description questionnaire: Comparison of factor structure in Australia, Spain and Turkey. *Research Quarterly for Exercise and Sport, 73,* 257–270.

Nideffer, R. (1993). Attention control training. In R.N. Singer, M. Murphey, & I.K. Tennant (Eds.), *Handbook of research on sport psychology* (pp. 127–170). New York: Macmillan.

Nisbett, R. & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Reardon, J.P. (1993). Handling the self-talk of athletes. In K.P. Henschen & W.F. Straub (Eds.), *Sport Psychology: An analysis of athlete behavior* (3rd ed., pp. 203–211). Forest Glen, MI: Mouvement Publications.

Royal, K.A., Farrow, D., Mujika, I., Halson, S.L., Pyne, D., & Abernethy, B. (2006). The effects of fatigue on decision making and shooting skill performance in water polo players. *Journal of Sports Sciences, 24,* 807–815.

Sarason, I.G., Sarason, B.R., Keefe, D.E., Hayes, B.E., & Shearin, E.N. (1986). Cognitive interference: Situational determinants and traitlike characteristics. *Journal of Personality and Social Psychology, 51,* 215–226.

Smith, R.E., Smoll, F.L., & Schutz, R.W. (1990). Measurement and correlates of sport-specific cognitive and somatic trait anxiety: The Sport Anxiety Scale. *Anxiety Research, 2,* 263–280.

Theodorakis, Y., Hatzigeorgiadis, A., & Chroni, S. (2008). The Functions of Self-Talk Questionnaire: Investigating how self-talk strategies operate. *Measurement in Physical Education and Exercise Science, 12,* 10–30.

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.

Van Raalte, J.L., Brewer, B.W., Rivera, P.M., & Petipas, A.J. (1994). The relationship between observable self-talk and competitive junior tennis players' performances. *Journal of Sport & Exercise Psychology, 16,* 400–415.

Vygotsky, L.S. (1986). In A. Kozulin (Ed.), *Thought and language*. Cambridge, Massachusetts: The MIT Press.

Weinberg, R.S., Gould, D., & Jackson, A. (1980). Cognition and motor performance: effect of matching-up strategies on three motor tasks. *Cognitive Therapy and Research, 4,* 239–245.

Zervas, Y., Stavrou, N.A., & Psychountaki, M. (2007). Development and validation of the Self-Talk Questionnaire (S-TQ) for Sports. *Journal of Applied Sport Psychology, 19,* 142–159.