Achievement-Motivated Behavior in Individual Sports: Evidence for the Construct and Criterion Validity of the AMBIS-I Coach-Rating Scale

Claudia Zuber  , Michael J. Schmid and Achim Conzelmann
Institute of Sport Science, University of Bern, Switzerland

Abstract
The coach-rating scale for Achievement-Motivated Behavior in Individual Sports (AMBIS-I) was constructed to measure achievement motivation, not from athletes’ own views but from coaches’ perspectives. The tool was already checked for reliability as well as content, factorial, and concurrent criterion validity (Zuber and Conzelmann, 2019). To further establish construct and criterion validity, two different samples were involved. Sample 1 included 67 experienced coaches rating their 278 athletes on the three AMBIS-I dimensions proactivity, ambition and commitment. In sample 2, 157 athletes completed self-report questionnaires measuring motivational and volitional concepts. Congruent and discriminant construct validity were assessed using the QCV-procedure (Westen and Rosenthal, 2003) by comparing experts’ predicted and empirically observed correlations between the coaches’ ratings on the AMBIS-I with the self-ratings of validated instruments. Consistent with theoretical expectations, achievement goal orientations, self-determination and self-optimization show significant positive relationships to the AMBIS-I scales, the negatively formulated volitional concepts, negative ones. As indicated by the $0.87 \leq \text{raters-CV} \leq 0.95$, the general patterns of the expert’s predictions triangulate consistently with the observed correlations. The findings concerning absolute agreement were mixed. Even though the ICCs suggest sufficient to good consistency, the values of $r_{\text{football-CV}}$ are considerably lower. To indicate criterion validity, AMBIS-I display medium to large correlations with the actual performance level estimated by the coaches and small to medium correlations with the assigned potential for subsequent success one year later. In summary, we found solid indications, that AMBIS-I is a valid measure of achievement-motivated behavior in individual sports from coaches’ perspective.

Key words: motivation, talent identification, talent development, individual sports, test construction.

Introduction
In talent research, the great importance of psychological variables for a successful development from a young promising to a successful top-level athlete is pointed out. Achievement motivation, in particular, seems to play a critical role for talent development and subsequent success (e.g. Abbott and Collins, 2004; Coetzee et al., 2006; Zuber et al., 2015). Motivational characteristics during training are also highly valued by coaches (Christensen, 2009; Jokuschies et al., 2017) as well as by elite athletes and their parents (MacNamara et al., 2010). However, when it comes to talent selection decisions, the assessment of achievement motivation is associated with significant concerns such as social desirability (see Zuber and Conzelmann, 2019, for a broader discussion of the methodological shortcomings of the different assessment methods). To remediate these methodological shortcomings, we decided to use the long experiences of coaches instead of coming back on the more subjective self-ratings of athletes and to focus on observable achievement-motivated behaviors as opposed to non-directly observable motives. Therefore “a suitable new tool for assessing achievement motivation in the context of selection decisions in sports should be based on coaches’ ratings of achievement-motivated behavior” (Zuber and Conzelmann, 2019, p. 3) and thus corresponds to the request to record psychological characteristics by means of observable behavior patterns (Musculus and Lobinger, 2018).

To design a reliable, valid and time-efficient instrument for thoroughly assessing achievement-motivated behaviors in sports, we built upon the act-frequency approach, which relies on the definitions of constructs elaborated by psychological laypersons (Buss and Craik, 1983). In a series of three studies, the final ten-item version of the AMBIS-I coach-rating scale was constructed and checked for reliability, as well as content, factorial and concurrent criterion validity. The three factors proactivity, ambition and commitment presented acceptable to good reliability and good factorial validity. In addition, relationships with athletes’ performance level point to the instrument’s concurrent criterion validity (Zuber and Conzelmann, 2019). As further evidence for construct and predictive criterion validity of AMBIS-I is still missing to date, the purpose of the current study aims to fill this gap.

Construct validity
Despite the rich theoretical body on achievement motivation, AMBIS-I is the first tool measuring achievement-motivated behavior in sports. So careful construct validation is of high relevance. Construct validity refers to the “degree to which test scores represent an individual’s standing on the theoretical construct the test is designed to measure” (Sireci and Sukin, 2013). That is often investigated by checking the “...match between a measure’s actual associations with other measures and the associations that the test should have with the other measures” (Furr and Bacharach, 2014).

The three subscales and ten items included in the AMBIS-I coach rating scale were not theory driven, but constructed inductively based on a prototype strategy. According to the requirements of the act-frequency approach (Buss and Craik, 1983), as the instrument was intended to be used by youth elite sports coaches, they were asked for...
relevant, observable behaviors that best capture the construct achievement motivation (Zuber and Conzelmann, 2019). It was found that instruments constructed with the prototype approach display comparatively good validity coefficients (Broughton, 1984). But how can the three factors of the AMBIS-I be contextualized in the nomological network surrounding the broadly examined construct of achievement motivation? In literature in personality psychology – which includes also motivation – the focus lays mainly on motives and not on situations or behaviors (Furr and Funder, 2010). There is no theory on achievement-motivated behavior in sports either. Therefore, to hypothesize on the position of achievement motivated behavior in the nomological network of achievement motivation, we have to refer to motivational constructs such as motives or achievement goal orientations, even though they don’t lie on the same level as behaviors. Rather, motives and goals build the foundation for behavior as it is deduced from the general model of determinants and course of motivated action (Heckhausen and Heckhausen, 2010). In the following section the three AMBIS-I factors are therefore connected with well examined motivational constructs in order to make the assumed relations empirically checkable.

“The factor proactivity refers to getting involved in training processes on one’s own initiative and for one’s own sake” (Zuber and Conzelmann, 2019, p. 8). The impulse to act is self-determined and does not require an external push, as when an athlete looks for opportunities to catch up on missed training content. Intrinsic motivation as the most self-determined motivation according to self-determination theory (Ryan and Deci, 2000) is characterized by the sensation of immanent pleasure associated with performing an activity for its own sake. On the opposite end of the self-determination continuum, extrinsic motivation pertains to actions triggered by expected consequences, such as fame, honor or prize money. Zuber and Conzelmann (2019) name also connections between proactivity and the concept of achievement goal orientations: A proactive athlete pursues his or her goals perseveringly, as does an athlete which is first on the training grounds and practicing technical processes independently. Achievement goal orientations guide actions towards certain goals, including task and ego orientation (Nicholls, 1984) or mastery and performance orientation (Ames and Archer, 1988), and competitiveness (Gill and Deeter, 1988). As the factor proactivity gives no indication of which goals are being closely pursued, all three discussed goal orientations are conceivable as the origin of proactive behavior. Due to the fact, that those goals are being pursued persistently, Zuber and Conzelmann (2019) assume also a connection to the concept of volition. Processes such as self-motivation and self-regulation are responsible for initiating an action and maintaining it until a specific goal has been achieved, what is especially important in the course of an athletic career (Elbe et al., 2005; Baron-Thiene and Alfermann, 2015). Therefore, starting from this theoretical positioning, it can be assumed that the factor proactivity has connections with self-determination, achievement goal orientations, as well as volitional processes.

“The factor ambition is characterized by the absolute will to successfully pursue self-imposed goals in competitions” (Zuber and Conzelmann, 2019, p. 8). Ambitious athletes aim at winning competitions which is for instance shown by athletes that clearly communicate before the competition that they want to win. If winning is not achieved, the athlete faces disappointment. Thus, the factor ambition shows conceptual similarities to the achievement goal orientations competitiveness and goal orientation (as the goals are self-imposed) and, probably to a smaller extent, ego orientation.

The factor commitment in the setting of training refers to readiness and willingness to perform (Zuber and Conzelmann, 2019). A committed athlete shows an active stance in training and demonstrates his/her will to work hard, to attain a goal, and to pursue a task repeatedly, even in the face of adversity (Scanlan et al., 1993), what again makes its connection with volition obvious (Elbe et al., 2005). Committed athletes also orient towards and compare with stronger athletes, what makes a connection with competitiveness realistic.

The theoretical embedding of the three factors just outlined yields certain assumptions on which already well-examined constructs are related to the achievement-motivated behavior measured with AMBIS-I. In summary, convergent validities of the scales of AMBIS-I with self-determination, achievement goal orientations, and volition are to be expected.

**Criterion validity**

Concurrent and predictive criterion validity are of particular importance in the context of talent identification and development because potential talent predictors and the tools to assess those talent characteristics should show associations with actual performance and must be able to predict later success or future performance (Buekers et al., 2015; Feichtinger and Höner, 2014; Morris, 2000; Sarmento et al., 2018; Vaeyens et al., 2008).

Previous studies have shown that motivational and volitional characteristics are relevant predictors of current (e.g. Dominguez-Escribano et al., 2017; Erikstad et al., 2018; Feichtinger and Höner, 2014; Gillet et al., 2012; Gonçalves et al., 2011) and future performance (e.g. Foman et al., 2016; Höner and Feichtinger, 2016; Unierzyński, 2003; Zuber and Conzelmann, 2014; Zuber et al., 2015) or dropout (e.g. Sarrazin et al., 2002). Motivational characteristics like achievement goal orientations (e.g. Dominguez-Escribano et al., 2017; Feichtinger and Höner, 2014; Gonçalves et al., 2011; Höner and Feichtinger, 2016; Zuber et al., 2015) and less common self-determination (e.g. Gillet et al., 2012; Zuber et al., 2015) as well as volitional characteristics (e.g. Erikstad et al., 2018; Feichtinger and Höner, 2014; Toering et al., 2009) displayed small to moderate associations to divers performance relevant criteria, including coach ratings, selection to a higher squad, professional contracts and objective performance in competitions. Many of these studies were conducted in team sports (mainly football) (e.g. Erikstad et al., 2018; Feichtinger and Höner, 2014; Foman et al., 2016; Gonçalves et al., 2011; Höner and Feichtinger, 2016; Toering et al., 2009; Zuber et al., 2015). Some studies refer to a mixed sample of team
and individual sports (Elbe and Beckmann, 2006; Halvari and Thomassen, 1997; Hellandsig, 1998; Mahoney et al., 1987; Raglin et al., 1990) and only a few examined only athletes from individual sports (Cervelló et al., 2007; Gillet et al., 2012; Unierzyski, 2003).

However, there is also contradictory evidence: Figueiredo and colleagues (2009) found no differences in achievement goal orientations between those athletes that dropped out or competed at club and elite level two years later. Additionally, there are results showing that elite athletes display lower levels of self-determined motivation than athletes on lower performance levels (Chantal et al., 1996; Fortier et al., 1995).

Nevertheless, as the vast majority of findings point in this direction, the achievement-motivated behavior recorded with AMBIS-I should display positive associations with current and future performance criteria.

The present research
To achieve our goal of examining the construct and criterion validity of the AMBIS-I coach-rating scale, we included the ratings of the coaches on the achievement-motivated behavior of their athletes as well as the self-ratings of the athletes on several related constructs. Hence, the consecutive analysis will be guided by the following three research questions. 1) What is the relationship between AMBIS-I and established self-report questionnaires capturing motivational and volitional constructs (construct validity) and 2) is AMBIS-I capable of predicting relevant performance criteria both cross-sectional (concurrent criterion validity and 3) longitudinally (predictive criterion validity)?

Methods

Participants and procedure
Formal ethical approval was granted from the authors’ institutional review board before conducting the study. The samples of coaches were recruited directly through the sport federations via Swiss Olympic. The sport directors of the sport federations categorized by Swiss Olympic in the categories 1 to 3 (of 5), according to their national importance and achievement potential, were asked to send lists of all their coaches at the 1st or 2nd level of education (professional training for elite or competitive sports). For this study, only German-speaking coaches were included. Using the lists provided by the sport federations, 160 coaches were invited to participate and 69 coaches completed the study at the first measurement point (t1), evaluating 288 athletes in individual sports with AMBIS-I (see Figure 1 as flow chart of the participants). In the pre-analysis, the assessments of two coaches were removed from the dataset, as they indicated, that they had known their athletes for less than half a year, or that they did not feel certain in their ratings. The final sample of 67 coaches (19 women, 28.4%, 48 men, 71.6%, \(M_{age} = 41.88, SD = 11.96\)) included 278 ratings from badminton, biathlon, curling,

![Figure 1. Participant flow chart.](image-url)
freestyle skiing, golf, judo, artistic cycling, cross country skiing, track and field, mountain biking, road cycling, sledding, rowing, swimming, alpine skiing, shooting, tennis, and equestrian vaulting. The coaches reported a high level of education, with more than 50% having successfully completed the highest or 2nd highest level of vocational training. They have $M = 16.19$ ($SD = 10.93$) years of professional experience and have known the athletes they evaluated on average for $M = 4.11$, $SD = 3.45$ years. To examine the stability of the results, $M = 3.2$, $SD = 0.72$ months after $t_1$, 46 coaches (16 women, 35%, 29 men, 65%, $Mage = 42.95$, $SD = 11.67$ years) participated and conducted ratings of 175 athletes for the second time ($t_2$).

The coaches involved in the investigation provided the contact details of 217 of their athletes who were subsequently invited to participate in the study (see Figure 1). All athletes received a cover letter along with a link to the online survey. Additionally, all parents of athletes younger than 16 years provided informed consent to record and use data for scientific research. The final sample of the athletes includes 157 athletes (62 women, 39.5%, 95 men, 60.5%; $Mage = 16.48$, $SD = 2.20$ years) practicing in their sport since $M = 6.55$, $SD = 2.96$ years. At the second measurement point ($t_2$), 114 athletes (73%) took part for a second time. Of 149 athletes ($t_1$: 81), the self- as well as the coach-rating were available. As 44 ($t_2$: 21) athletes of them were rated by their main as well as by their assistant coach, those self-ratings were used twice and matched once with the rating of the main coach and once with the rating of the assistant coach. Finally, 193 data sets consisting of a self- and a coach-rating were available for answering research question 1 ($t_1$: 102). To answer research questions 2 and 3, all 278 coach-ratings ($t_2$: 175) were integrated in the analysis.

All data were collected using an internet-based questionnaire (LimeSurvey, Version 2.50). Further information on the study design is available in Figure 2.

**Measures**

*Achievement-motivated behavior (coach-rating).* AMBIS-I consists of 10 prototypical behaviors whose occurrences were to be evaluated on a 4-point scale ranging from 0 (= never) to 3 (= often) with a “not able to respond” option. For the second measurement time, the answer format was extended by the category 4 (= always) (Zuber and Conzelmann, 2019). The items built the three factors proactivity (“He/she stayed after training to continue practicing”), ambition (“He/she has shown that he/she is not satisfied with 2nd place”) and commitment (“In high demanding exercises, he/she worked until exhaustion”). As the factors are intercorrelated, the total score of all items, was used as well. The high number of items offered more reliable statements, whereas the individual factors provided more detailed information (Zuber and Conzelmann, 2019).

As each coach had submitted a list with the names of athletes whom they were training at that moment, we were able to request their rating for each athlete individually: “How often did athlete A [name of one of the coach’s athletes] display the behavior mentioned below over the last 12 months?”

In addition, the coaches were asked how certain they felt in their assessment of the respective athlete (not at all, a little, somewhat, fairly much) and how long (in years) they had already known their respective athlete. Finally, the educational level of the coaches was assessed.

*Motivational and volitional constructs (self-rating).* To assess the convergent (and discriminant) construct validity of AMBIS-I (research question 1), all athletes were requested to fill in a series of scientifically established self-report questionnaires designed to capture motivational and volitional constructs. As especially convergent construct validity of AMBIS-I to self-determination, achievement goal orientations, and volition seem to be relevant (as discussed above), scales measuring those three concepts were included (Table 1). All used scales displayed reasonable to good internal consistencies at $t_1$ (Table 1).

*Performance criteria.* As an external performance criterion for answering research question 2, all coaches were asked to rate the performance level of the athletes on an 8-point scale at the same time as they completed the AMBIS-I ratings (cross-sectional) with 0 (very low level) to 7 (international level). As a second criterion and to answer research question 3, we checked the Swiss Olympic Cards (SOC) of the participating athletes one year after $t_1$. **Figure 2. Study design.**
Data analysis

Data were analyzed using IBM SPSS Statistics 24. Overall, the amount of missing data lies – depending on sample and measurement point – between 0.6% and 3.7% and was imputed using the expectation-maximization algorithm. For examining the construct validity (research question 1), the QCV (quantifying construct validity)-procedure was used (Westen and Rosenthal, 2003), in which the fit between the theoretically derived predictions of correlations is compared with the empirically found correlations. This allows to avoid that the determination of convergent and discriminant construct validity leads to an overly subjective interpretation of the obtained set of correlations (Furr and Bacharach, 2014). To determine which predictions were to be assumed theoretically, the procedure described by Furr and Bacharach (2014) was used and five experts in the field were recruited. They were all working at academic institutions for sport science or psychology and were experienced in publishing and teaching on the topic of motivation in sports. The raters were given brief descriptions of the scales shown in Table 2 and were then asked to provide a point estimate to one decimal place, of their predicted correlation. To check whether the experts agreed in their assumptions in principle, inter-rater reliability was first computed using two-way random effects, absolute agreement intraclass correlations (ICCs; Koo and Li, 2016). As these values displayed – according to the recommendations of Koo and Li (2016) – good inter-rater reliability (see Table 3), the five sets of predictions were then averaged in one single set of predictions. Then, the two correlation coefficients ralerting-CV and rcontract-CV for quantifying construct validity were computed. According to Westen and Rosenthal (2003) both coefficients “provide a quantified index of the degree of convergence between the observed pattern of correlations and the theoretically predicted pattern of correlations – that is the degree of agreement of the data with the theory underlying the construct and the measure” (p. 608). The coefficient ralerting-CV represents the correlation between the patterns of the predicted and the observed correlations and is therefore easy computable and interpretable. Whereas rcontract-CV is seen as a “rough, readily interpretable index that can alert the researcher to possible trends of interest” (Westen and Rosenthal, 2003), rcontract-CV is a more strict test of congruence between expected and obtained associations. It is derived from a series of complex calculations (see Westen and Rosenthal, 2003) and takes into account the median intercorrelations, the absolute values of the observed correlations, and the sample size. This overall index can also be interpreted as a correlation coefficient and tested for statistical significance. As there were no benchmarks for evaluating the magnitude of ralerting-CV and rcontract-CV and the method is not widely used yet, we used the ICC as additional index.

The ICC assesses the consistency between judges’ ratings of a set of objects (Field, 2009) and therefore two sets of correlations (observed and predicted correlations) were treated as the estimates of two different raters. The used two-way random effects, absolute agreement intraclass correlations (ICC) take the relative differences between raters as an important source of disagreement (Field, 2005) and is thus a stricter test than ralerting-CV.

The interpretation of the ICCs is based on the recommendations of Koo and Li (2016), who classified ICCs lower than .50 as poor, between 0.50 to 0.74 as sufficient, between 0.75 to 0.89 as good and higher than 0.90 as excellent.

To assess criterion validity (research questions 2 and 3), Pearson correlations between AMBIS-I coach ratings and estimated performance levels at the same measurement time (concurrent criterion validity) and Swiss Olympic Cards one year later (predictive criterion validity) were used. The interpretation of the effect sizes of the correlations was calculated according to the recommendations of Gignac and Szodorai (2016) with r = 0.10, 0.20, and 0.30 as small, medium, and large effects, respectively.

Table 1. Self-report instruments used for the construct validation of AMBIS-I.

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Construct</th>
<th>Scale</th>
<th>α</th>
<th># Items</th>
<th>Sample item (translation from German original items)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sport Orientation Questionnaire (SOQ)</td>
<td>Achievement goal orientation</td>
<td>Competitiveness (CO)</td>
<td>.87</td>
<td>13</td>
<td>I’m looking forward to competitions.</td>
</tr>
<tr>
<td>(Elbe, 2004)</td>
<td></td>
<td>Win Orientation (WO)</td>
<td>.83</td>
<td>6</td>
<td>I have the most fun when I win.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Goal Orientation (GO)</td>
<td>.79</td>
<td>6</td>
<td>I try hardest when I have a specific goal.</td>
</tr>
<tr>
<td>Sport Motivation Scale (SMS)</td>
<td>Self-determination</td>
<td>Se (SDI)</td>
<td>.83</td>
<td>28</td>
<td>It gives me pleasure to learn more about my sport.</td>
</tr>
<tr>
<td>(Burtscher et al., 2011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volitional Components in Sport (VCS)</td>
<td></td>
<td>Self-Optimization (SO)</td>
<td>.93</td>
<td>29</td>
<td>I am optimistic about most things in sports.</td>
</tr>
<tr>
<td>(Wenhold, Elbe &amp; Beckmann, 2008)</td>
<td></td>
<td>Self-Impediment (SI)</td>
<td>.78</td>
<td>9</td>
<td>When I get into a bad mood during sports, I can hardly get out of it.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of Initiation (LI)</td>
<td>.88</td>
<td>13</td>
<td>I usually only start to train properly when someone puts pressure on me.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Loss of Focus (LF)</td>
<td>.78</td>
<td>9</td>
<td>In training, I often have to think about things that have nothing to do with what I’m doing.</td>
</tr>
</tbody>
</table>

*The seven subscales Intrinsic Motivation “To Know”, “To Accomplish” and “To Experience”, External, Introjected and Identified Regulation, as well as Amotivation were combined to form a Self-Determination Index (see Vallerand, 2001)
Results

As shown by the descriptive statistics in Table 2, proactivity was least likely to be displayed, followed by ambition and commitment. Additionally, commitment exhibits restricted variance, especially at t1. The higher values in all dimensions at t2 compared with t1 were at least partly due to the change in the scaling from 0–3 at t1 to 0–4 at t2.

Construct validity

To assess congruent and discriminant construct validity of AMBIS-I, a comparison of a theoretically derived set of correlations with the empirically obtained correlations between the coach-ratings on AMBIS-I and the self-ratings at t1 on the already validated instruments AMS, SOQ, SMS and VQS, that capture motivational and volitional constructs, was conducted. These actual observed and the predicted correlations are displayed in Table 3. All significant observed correlations lie within a low to medium range, with $r = 0.33$ being the highest value for the correlation between proactivity and competitiveness.

The majority of the resulting correlations were in line with the expectations deducted from the nomological network of constructs. Proactivity displayed the highest numbers of relations with the self-report on motivational and volitional constructs. Athletes described as being high in proactivity by their coach, tend to score higher in competitiveness, win and goal orientation, self-determination and self-optimization, and lower in loss of focus. These athletes liked to compete and to win, pursued their own goals, acted in their sport for self-determined reasons and developed strategies and will processes that support the pursuit of goals and the implementation of actions. As expected, ambitious athletes showed the highest positive correlations with the achievement goal orientations competitiveness and goal orientation and a negative correlation coefficient with the factor loss of focus, which is seen as a volitional deficit. As theoretically assumed, committed athletes then, displayed higher values for competitiveness and again lower values for loss of focus. The total score displayed in many cases higher values than the single scales and therefore a satisfactory construct validity.

To examine the stability of the results, the resulting observed correlations at t2 are displayed in Table 4. The correlations of t1 and t2 are by and large in a similar area. Two not expected exceptions should be mentioned. First, there was no correlation of self-determination with the total score and proactivity. Second, the relations with the volitional factors somehow changed, but still pointed in the expected direction, as there were higher correlations with the factor self-optimization, but lower ones with loss of focus at t2.

All in all, the experts overestimated the degree to which the coach-ratings of the factors would correlate with the self-report questionnaires. Whereas the values of the ICCobs-pred and the Faltering-CV were good for all three factors, the Fcontrast-CVs were comparatively low (Table 3). As displayed with Faltering-CV, the general patterns of the expert’s predictions were consistent with the observed correlations. The findings to the absolute agreement were somewhat ambiguous: Whereas the ICCobs-pred points to sufficient to good consistence, the values of Fcontrast-CV were considerably lower.

It can be summarized, that the results largely reflected the expectations of the experts and theoretical considerations and for the most part, showed stability. In the magnitude of the correlations, however, cuts must be made.

Table 2. Descriptive statistics and Cronbach’s $\alpha$ of the AMBIS-I factors and the total score at t1 and t2.

<table>
<thead>
<tr>
<th>Scales</th>
<th>Proactivity</th>
<th></th>
<th>Ambition</th>
<th></th>
<th>Commitment</th>
<th></th>
<th>Total Score AMBIS-I</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t1 (n = 278)</td>
<td>t2 (n = 175)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>$\alpha$</td>
<td>M</td>
<td>SD</td>
<td>$\alpha$</td>
<td></td>
</tr>
<tr>
<td>Proactivity</td>
<td>1.51</td>
<td>0.73</td>
<td>.81</td>
<td>1.76</td>
<td>0.89</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td>Ambition</td>
<td>1.81</td>
<td>0.77</td>
<td>.78</td>
<td>2.25</td>
<td>0.94</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td>Commitment</td>
<td>2.40</td>
<td>0.53</td>
<td>.64</td>
<td>2.80</td>
<td>0.67</td>
<td>.71</td>
<td></td>
</tr>
<tr>
<td>Total Score AMBIS-I</td>
<td>1.91</td>
<td>0.52</td>
<td>.82</td>
<td>2.27</td>
<td>0.69</td>
<td>.89</td>
<td></td>
</tr>
</tbody>
</table>

* t1: Scale 0-3; t2: Scale 0-4

Table 3. ICCs, Faltering-CV and Fcontrast-CV to define construct validity patterns of observed and predicted correlations between coach-ratings of AMBIS-I and the self-report questionnaires at t1 (n = 193).

<table>
<thead>
<tr>
<th>Scales</th>
<th>Proactivity</th>
<th></th>
<th>Ambition</th>
<th></th>
<th>Commitment</th>
<th></th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>Predicted</td>
<td>Observed</td>
<td>Predicted</td>
<td>Observed</td>
<td>Predicted</td>
<td>Observed</td>
</tr>
<tr>
<td>SOQ CO</td>
<td>.33*</td>
<td>.18</td>
<td>.21*</td>
<td>.38</td>
<td>.24*</td>
<td>.28</td>
<td>.34*</td>
</tr>
<tr>
<td>SOQ WO</td>
<td>.22*</td>
<td>.12</td>
<td>.13</td>
<td>.44</td>
<td>.12</td>
<td>.16</td>
<td>.21*</td>
</tr>
<tr>
<td>SOQ GO</td>
<td>.25*</td>
<td>.34</td>
<td>.15*</td>
<td>.32</td>
<td>.13</td>
<td>.30</td>
<td>.23*</td>
</tr>
<tr>
<td>SMS SDI</td>
<td>.22*</td>
<td>.38</td>
<td>.02</td>
<td>.08</td>
<td>.12</td>
<td>.26</td>
<td>.15*</td>
</tr>
<tr>
<td>VQS SO</td>
<td>.25*</td>
<td>.36</td>
<td>.11</td>
<td>.14</td>
<td>.14</td>
<td>.24</td>
<td>.22*</td>
</tr>
<tr>
<td>VQS SI</td>
<td>-.02</td>
<td>-.10</td>
<td>.12</td>
<td>.10</td>
<td>-.07</td>
<td>-.12</td>
<td>-.03</td>
</tr>
<tr>
<td>VQS LI</td>
<td>-.14</td>
<td>-.38</td>
<td>-.06</td>
<td>-.18</td>
<td>-.12</td>
<td>-.34</td>
<td>-.13</td>
</tr>
<tr>
<td>VQS LF</td>
<td>-.22*</td>
<td>-.16</td>
<td>-.15*</td>
<td>-.20</td>
<td>-.16*</td>
<td>-.26</td>
<td>-.23*</td>
</tr>
<tr>
<td>ICC Rater$^b$</td>
<td>.96*</td>
<td>.97*</td>
<td></td>
<td>.94*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICC obs-pred$^b$</td>
<td>.85*</td>
<td>.70*</td>
<td></td>
<td>.82*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faltering-CV</td>
<td>.87*</td>
<td>.90*</td>
<td></td>
<td>.95*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fcontrast-CV</td>
<td>.45*</td>
<td>.29*</td>
<td></td>
<td>.35*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $p < 0.05$ (two-sided); CO = competitiveness; WO = win orientation; GO = goal orientation; SDI = self-determination index; SO = self-optimization; SI = self-impediment; LI = lack of initiation; LF = loss of focus. $^a$ mean of 5 raters; $^b$ between the observed and the predicted correlations.
Criterion validity

All three AMBIS-I scales displayed significant correlations with the actual performance level as rated by their coaches with medium to large effects (Table 5). The total score showed even higher correlations and therefore suggested reasonable concurrent criterion validity. The values for predictive criterion validity lied slightly lower but display along with the exception of commitment at t2 still significant effects.

Discussion

The current study aimed to assess the construct and criterion validity of the newly developed AMBIS-I coach-rating scale designed to measure achievement-motivated behavior in individual sports. AMBIS-I assessed the three factors proactivity, ambition, and commitment, based on ten forms of behavior, whose frequency of occurrence is assessed by the coaches.

Overall, the results on construct validity were evaluated positively: Consistent with theoretical expectations, achievement goal orientations, self-determination, and self-optimization suggested significant positive, the volitional deficits negative relationships to all three or single AMBIS-I scales and the total score. As displayed with high rcontrast-CVs, the general patterns of the expert’s predictions matched with the observed correlations. The findings for the absolute agreement were mixed: Whereas the ICCs point to sufficient to good consistency, the values of ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

Two reasons could have contributed to the rather low correlations between the coach- and the self-ratings. First, the constructs used for validation (motives) don’t lie on the same level as the behaviors. And as it is assumed that there is “no direct link between a specific behavior and a specific personality characteristic” (Furr and Funder, 1997). So, the reliability coefficients for evaluating the magnitude of rcontrast-CVs (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

Two reasons could have contributed to the rather low correlations between the coach- and the self-ratings. First, the constructs used for validation (motives) don’t lie on the same level as the behaviors. And as it is assumed that there is “no direct link between a specific behavior and a specific personality characteristic” (Furr and Funder, 1997). So, the reliability coefficients for evaluating the magnitude of rcontrast-CVs (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

Two reasons could have contributed to the rather low correlations between the coach- and the self-ratings. First, the constructs used for validation (motives) don’t lie on the same level as the behaviors. And as it is assumed that there is “no direct link between a specific behavior and a specific personality characteristic” (Furr and Funder, 1997). So, the reliability coefficients for evaluating the magnitude of rcontrast-CVs (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.

that there are no benchmarks for evaluating the magnitude of rcontrast-CVs coefficients (Thomas et al., 2012). Because few studies have been conducted using the Westen and Rosenthal (2003) method to test the validity of psychological measures, it is difficult to judge rcontrast-CVs coefficients. Accordingly, we computed an additional coefficient, the ICC, which was assumed to be an appropriate measure to assess the consistency between the observed and predicted correlation patterns, and which displays satisfactory agreement.
From a content point of view, it is noticeable that the three factors of AMBIS-I can hardly be separated from each other by the combination of the relationships to other motivational and volitional constructs. Only proactivity can be distinguished with higher correlations to self-determination and self-optimization at t1. All three factors are positively related to competitiveness and negatively related to the volitional deficit loss of focus. Ambition can also score by a positive connection to goal orientation. However, this proximity in terms of content of the three factors is not unexpected. After all, the three factors have all been constructed under the construct of achievement-motivated behavior (see Zuber and Conzelmann, 2019). However, the different frequencies of occurrence of different behaviors can still provide valuable information for talent development. It would also be interesting to examine, from a person-oriented approach, whether different combinations of the three factors lead to different outcomes (e.g. Zuber et al., 2015). The quality and intensity of motivation was found to determine the adaptive or maladaptive quality of performance outcomes (Lemyre et al., 2007). It should therefore be tested, whether the highest possible values in achievement-motivated behavior must be rated positively in every case, or whether extreme manifestations, for example in proactivity or commitment, are at risk to be accompanied by negative phenomena such as over-involvement or burnout (Gardner and Moore, 2006). For example, there was found that highly committed athletes may feel entrapped in their sport behavior and thus a link between burnout and the character and level of commitment might be existent (Gustafsson et al., 2011).

To summarize, there are for the most part evidence for accurate construct validity, although the different methods of self- and coach-rating make the interpretation somewhat more difficult. In particular, the overall score, however, shows correlations in direction and level that would have been expected from achievement-motivated behavior towards explicit goals and volition. Therefore, achievement-motivated behavior in sport can reasonably be contextualized in the nomological network surrounding the broadly examined construct of achievement motivation, and can be defined as self-determined behavior in the context of competitive sports, which aims to achieve competition- or task-oriented goals and which involves a high degree of self-regulation and commitment. The results for concurrent and predictive criterion validity turn out as expected. Whereas the concurrent validity uncovers medium to large correlations with the coach-ratings of performance, the prediction of the performance level one year later marked by the Swiss Olympic Cards, demonstrated small to medium effects. There may be two reasons for this decline in predictive power. First and fairly obvious, the prediction over a longer period is a stricter test. Second, by using the SOCs as criteria, we tried to diminish possible bias of coaches, as in most cases the decision to award a SOC is not made by the own coach, but is rather the result of a combination of assessments of different national coaches and objective test results. However, the prognostic period of one year is rather short, especially if we remember that talent tools should predict performance at the age peak performance. For this reason, the prognostic values of AMBIS-I need reexamination again in the upcoming years.

Furthermore, the different methods for assessing construct validity by comparison of the coach-ratings of AMBIS-I and self-rating instruments as validation tools, has – as we have seen – disadvantages. As to the best of our knowledge, there is no comparable tool to assess achievement-motivated behavior from the coaches’ perspectives, hence no other option was available. In addition to extending the time period for predicting the performance level (see above), it would also be beneficial to include objective performance criteria (e.g. results in competitions or performance tests). Future research should furthermore provide reference values obtained by a larger and representative sample and check whether this or an adapted version of AMBIS would be a reliable and valid tool for assessing achievement-motivated behavior in team sports.

**Conclusion**

Overall, it seems that achievement-motivated behavior is a construct, which is linked with performance outcomes whereas AMBIS-I has proven to be a reliable and valid coach-rating scale. It could therefore – along with instruments to assess other performance relevant areas (e.g. for soccer: Williams and Franks, 1998) – be a noteworthy tool to be used in a multidimensional test battery for talent identification. In addition, because talent studies in the field of individual sports are so far rarely conducted, the present study provides important insights in the field of talent selection and talent promotion in these sports.

**Acknowledgements**

We thank Swiss Olympic and the Swiss Federal Office for Sport for supporting and funding this research project. Additionally, we wish to acknowledge the help provided by Nina Schorno for her assistance in the data collection and Olivier Schmid for proofreading an earlier version of this manuscript and giving helpful comments. The experiments comply with the current laws of the country in which they were performed. The authors have no conflict of interest to declare.

**References**


Key points

- The coach-rating scale AMBIS-I can be used for the reliable and valid assessment of achievement motivated behavior in individual sports from coaches’ perspectives for talent identification and promotion.
- We strongly recommend using AMBIS-I as part of a multidimensional test battery for talent identification.
- Low agreement between coach and athlete assessment is not surprising and should be discussed and used for further goal agreement.