

Ger J Exerc Sport Res
<https://doi.org/10.1007/s12662-019-00600-6>
 Received: 5 July 2018
 Accepted: 12 June 2019

© The Author(s) 2019



Claudia Zuber · Achim Conzelmann

Universität Bern, Bern, Switzerland

Achievement-Motivated Behavior in Individual Sports (AMBIS-I)—Coach Rating Scale

Development and preliminary validation

Electronic supplementary material

The online version of this article (<https://doi.org/10.1007/s12662-019-00600-6>) contains supplementary material, which is available to authorized users.

Talent research highlights the great importance of psychological variables for the successful development from a promising young to a successful top athlete (Coetzee, Grobbelaar, & Gird, 2006; Johnston, Wattie, Schorer, & Baker, 2018; MacNamara, Button, & Collins, 2010). Achievement motivation, in particular, seems to play a critical role for talent development and subsequent success (e.g. Abbott & Collins, 2004; Coetzee et al., 2006; Zuber, Zibung, & Conzelmann, 2015). Unlike the different theories subsumed under the umbrella term achievement motivation, achievement motivation itself is hardly defined precisely (Elliot & Dweck, 2007). Elliot and Dweck (2007) therefore propose to put the construct *competence*, “defined as a condition or quality of effectiveness, ability, sufficiency, or success” (Elliot & Dweck, 2007, p. 5), at the core of the achievement motivation literature.

Nevertheless, theory building and research on achievement motivation has had a huge impact in the last decades. The theories most currently discussed in talent research in sport are the achievement motives hope for success and fear of failure (e.g. Elbe & Beckmann, 2006; Halvari & Thomassen, 1997; Höner & Feichtinger, 2016; Sagar, Busch, & Jowett,

2010; Zuber et al., 2015; Zuber & Conzelmann, 2014), achievement goal orientations (e.g. Figueiredo, Gonçalves, Coelho e Silva, & Malina, 2009; Helmlandsig, 1998; Höner & Feichtinger, 2016; Huijgen, Elferink-Gemser, Lemmink, & Visscher, 2014; Reilly, Williams, Nevill, & Franks, 2000; Zuber et al., 2015) and the Self-Determination Theory (e.g. Gillet, Berjot, Vallerand, Amoura, & Rosnet, 2012; Gillet, Vallerand, & Paty, 2013; Gillet, Vallerand, & Rosnet, 2009; Sarrazin, Vallerand, Guillet, Pelletier, & Cury, 2002; Zuber et al., 2015).

The achievement motive determines whether individuals tend to approach achievement-related situations or whether they tend to avoid them (Atkinson, 1957). The positive connection between hope for success and performance in sports has been empirically confirmed in both cross-sectional (Coetzee et al., 2006; Halvari & Thomassen, 1997) and longitudinal studies (Elbe & Beckmann, 2006; Murr, Feichtinger, Larkin, O'Connor, & Höner, 2018; Unierzyski, 2003; Zuber & Conzelmann, 2014). Fear of failure, on the other hand, is often associated with a negative correlation with performance (Halvari & Thomassen, 1997; Sagar et al., 2010).

Whereas the achievement motive initiates actions aimed at attaining competence, achievement goal orientations guide these actions towards certain goals. It is differentiated between two or three different goal orientations, which are either called task and ego orientation (Nicholls, 1984) or mastery and performance orientation (Ames &

Archer, 1988). Task/mastery orientation is aimed at improving one's own skills, for which purpose an internal standard of comparison is used. Ego/performance orientation, on the other hand, focuses on displaying one's own superiority to other people (Heckhausen & Heckhausen, 2010). The third goal orientation competitiveness is characterized by “the desire to enter and strive for success in sport competition” (Gill & Deeter, 1988, p. 200). Therefore, it goes along with a need to compete and compare oneself in a sporting competition. The concept of achievement goals has fertilized a very large number of studies in sport sciences (Duda, 2007). Regarding the question on the relation between achievement goal orientations and achievement within youth elite sports, the major findings by the most recent review (in soccer) exposed a mixed state of evidence for ego/performance orientation with future performance. The majority of studies considering task/mastery orientation found a positive relation with future success (Murr et al., 2018).

In Self-Determination Theory, the reasons for motivated actions are distinguished according to where their perceived locus of causality is, or to what extent they are self-determined (Deci & Ryan, 1985). The resulting motivational type lies on a continuum extending from amotivation, a state with a complete absence of any motivation, through extrinsic motivation, and ending with intrinsic motivation as the most self-determined form of motivation (Ryan & Deci, 2000). Intrinsic motivation is char-

acterised by pleasure in performing the activity itself. Extrinsic motivation, on the other hand, pertains to actions which are carried out because of the expected consequences, such as fame, honour or prize money. Four types of extrinsic motivation are postulated, which are characterised by increasingly high levels of self-determination or autonomy (for an overview, see Ryan & Deci, 2007). A high degree of self-determination has been shown to be associated with higher levels of performance in adolescents (Gillet et al., 2009; Gillet et al., 2012; Zuber et al., 2015). Conversely, low levels of self-determination appears to hamper a successful sports career in the sense of dropping out (Calvo, Cervelló, Jiménez, Iglesias, & Murcia, 2010; Pelletier, Fortier, Vallerand, & Brière, 2001; Sarrazin et al., 2002).

As a result of (a) the empirical research pointing to the positive association between achievement motivation and performance, and (b), the experiences from sporting practice supporting that motivational characteristics are highly valued by coaches (Christensen, 2009; Jokuschies, Gut, & Conzelmann, 2017), as well as by elite athletes and their parents (MacNamara et al., 2010), the assessment of achievement motivation was integrated in multidimensional talent identification programs (TID; cf. Vaeyens, Lenoir, Williams, & Philippaerts, 2008; Germany: Feichtinger & Höner, 2014; Switzerland: Fuchslocher, Romann, Rüdisüli, Birrer, & Hollenstein, 2011). For the field of motor diagnostics, tests are generally regarded to provide a high degree of psychometric quality and usability (Zibung, Zuber, & Conzelmann, 2016), whereas the assessment of psychological variables, like achievement motivation for real selection decisions in contrast to scientific research, is associated with some serious problems, which will be described below.

To date, different methods to assess achievement motivation are in use. One can distinguish between (1) self-report questionnaires measuring the explicit achievement motive, (2) external assessment questionnaires (coach-ratings) measuring the explicit achievement mo-

tive, and (3) projective tests of the implicit achievement motive.

Self-report questionnaires—as the first method discussed here—became increasingly more available and were also used in sport science (Feichtinger & Höner, 2014) due to the high reliability and economical nature of self-reports. Questionnaires that have often been used in the context of achievement motivation in competitive sports are the Achievement Motivation Scale in Sport (Elbe, Wenhöf, & Müller, 2005b) to assess hope for success and fear of failure, the Sport Orientation Questionnaire (Elbe, 2004; Gill & Deeter, 1988) that focuses on the achievement goal orientations, and the Sport Motivation Scale (Burtscher, Furtner, Sachse, & Burtscher, 2011; Pelletier et al., 1995; Pelletier, Rocchi, Vallerand, Deci, & Ryan, 2013) to measure self-determination. In addition to the advantages that have led to the frequent use of self-report questionnaires, disadvantages are also known: “Although this approach has clear advantages, such as its high psychometric quality and ease of analysis, it also has its disadvantages. Responses may be biased by the tendency to present oneself in a socially desirable light ...” (Brunstein & Heckhausen, 2010, p. 138). The more obvious the purpose of the questionnaire, the more likely it is that the subject may intentionally or unintentionally influence the outcome, both in a positive and in a negative direction (response bias). Particularly in personally significant situations, the risk of socially desirable responding appears to increase: “Socially desirable responding might be more likely to occur in contexts in which important consequences hinge on the testing outcomes” (Furr & Bacharach, 2014, p. 280). In the context of TID, in which the athletes wish to be selected for the higher squad, the tendency to give socially desirable answers becomes particularly important. The athlete forms an opinion as to which answers or which characteristic expressions might have a positive influence on the selection, adjusting his/her answers in the corresponding direction. Assuring anonymity is recommended as one strategy for preventing or minimizing the bias of socially desirable responses

(Furr & Bacharach, 2014). This is possible by using codes or pseudonyms for research purposes. Thus, it has been shown that under anonymous conditions, youth footballers tended to score lower on positively associated personality traits than players under personalized conditions (Feichtinger & Höner, 2014). However, as anonymity cannot be ensured in selection-relevant contexts, the use of self-assessment questionnaires for selection decisions should be viewed very critically.

As this problem of socially desirable answers in self-rating questionnaires is well known, the use of external assessments or a coach's rating—as the second method discussed here—is recommended in amendment to the self-report questionnaires (Anshel & Lidor, 2009). However, this is associated with further methodological problems as the external raters have to evaluate statements that are not directly accessible to them. Motives have a low visibility as they often refer to internal emotions and thoughts and can therefore prove difficult to observe (Connelly & Ones, 2010; Furr & Funder, 2010). Consequently they have to be assessed with different indicators and using evidence from previous behavior as an aid, both of which are associated with the risk of memory effects (Buss & Craik, 1983). Accordingly, questionnaires on the external assessment of the achievement motive—such as those used by the Swiss National Olympic Committee (Swiss Olympic; Fuchslocher et al., 2011)—should only be used with great caution in TID.

As a third method, the assessment of the implicit achievement motive is discussed. The implicit achievement motive is not directly accessible and hence less affected by social desirability. For the assessment, projective methods, such as the Picture Story Exercise (PSE; Schultheiss & Pang, 2010) were used. However, assessment tools like these are time-consuming and uneconomical, and therefore cannot be recommended in TID.

Neither the explicit nor the implicit achievement motive seems to be assessable problem-free in the context of selection decisions using up-to-date diag-

nostic tools. Nevertheless, we must emphasize that the tools mentioned are not without use per se, but that these assessment difficulties arise specifically to the context of TID in sport. Instead of assessing motives, it is also conceivable to rely on recording concrete behavior, which, to put it very simplistically, results from the interaction between persons (e.g., motives) and situations (Heckhausen & Heckhausen, 2010). The advantage is that behavior can also be observed from the outside (e.g. by the coaches), and does not need to be investigated invasively. It is argued by Johnson (1997) that observers are better judges of behavior, whereas for more covert cognitions and feelings, self-reports are most accurate. Likewise, it can be deduced from the general model of determinants and course of motivated action (Heckhausen & Heckhausen, 2010, p. 3) that achievement-motivated behavior predicts future behavior and thus later performance better than a simple motive assessed by self-report that ignores situational factors. Besides, it seems reasonable to focus on decisions from coaches as they often form the basis for selection decisions in sport (Christensen, 2009). Coaches decide which athletes should make the leap to the next highest squad, where they will receive optimal support. For this decision, expert coaches can draw on many years of experience, and thus possess different opportunities for the comparison of different athletes: “Expert coaches can base their assessments of psychological characteristics on a representative sample of talented players they have worked with in the past ... [and] are able to make inter-individual comparisons and use this knowledge to evaluate and predict a player’s current and future potential” (Musculus & Lobinger, 2018, p. 2). In this way, talent characteristics identified by coaches seem to be relevant for the prognosis of future achievements (Jokuschies et al., 2017). Therefore, a suitable new tool for assessing achievement motivation in the context of selection decisions in sports should consequentially be based on coaches’ ratings of achievement-motivated behavior.

But what overt behaviors can one think of in terms of achievement-moti-

Ger J Exerc Sport Res <https://doi.org/10.1007/s12662-019-00600-6>
© The Author(s) 2019

C. Zuber · A. Conzelmann

Achievement-Motivated Behavior in Individual Sports (AMBIS-I)—Coach Rating Scale. Development and preliminary validation

Abstract

The assessment of achievement motivation in the context of selection decisions in elite youth sports is associated with serious problems (e.g. socially desirable responses). In order to counteract such problems, an external rating scale for the assessment of the Achievement-Motivated Behavior in Individual Sports (AMBIS-I) from a coach’s perspective was constructed and checked for psychometric quality in three consecutive steps. The studies are based on four different German-speaking samples, including 101 experienced coaches from individual sports and 26 sport psychologists. Multiple phases of exploratory structural equation modeling, item removal and cross-validation unveiled

a three-factorial model with 10 items displaying excellent fit indices, acceptable to good reliability, and evidence based on internal structure. Relationships with athletes’ performance level point to the instrument’s evidence for test-criterion relationship. These preliminary results are promising when considering the construction and show the potential of the economical coach rating scale for the scientifically sound assessment of observable achievement-motivated behavior in individual sports.

Keywords

Coach rating scale · Youth sports · Talent identification · Test construction · Psychology

Trainerbeurteilungsskala zum leistungsmotivierten Verhalten in Individualsportarten. Entwicklung und erste Validierung

Zusammenfassung

Die Diagnose der Leistungsmotivation im Kontext von Selektionsentscheidungen im Nachwuchsleistungssport ist mit ernsthaften Problemen assoziiert (z. B. sozial erwünschtes Antwortverhalten). Um diesen Problemen zu begegnen, wurde in 3 aufeinanderfolgenden Studien ein Fremdbeurteilungsinstrument für die Erfassung von leistungsmotiviertem Verhalten in Individualsportarten aus Trainersicht (AMBIS-I) konstruiert und bezüglich seiner Testgütekriterien überprüft. Die Studien basieren auf 4 unterschiedlichen deutschsprachigen Stichproben mit 101 erfahrenen Trainer(inne)n aus Individualsportarten sowie 26 Sportpsycholog(inn)en. Nach mehreren Phasen von explorativer Strukturgleichungsmodellierung, Item-elimination und Kreuzvalidierung ergab sich ein 3-Faktoren-Modell, welches aus

10 Items besteht und eine hervorragende Anpassungsgüte sowie akzeptable bis gute Reliabilität und faktorielle Validität aufweist. Die gefundenen Zusammenhänge mit dem Leistungsniveau der Athlet(inn)en deuten zudem auf die Kriteriumsvalidität des Instruments hin. Diese ersten Ergebnisse sind im Hinblick auf die weitere Validierung vielversprechend und zeigen das Potenzial dieser ökonomischen Trainerbeurteilungsskala zur wissenschaftlich fundierten Erfassung des beobachtbaren leistungsmotivierten Verhaltens in Individualsportarten auf.

Schlüsselwörter

Trainerbeurteilungsskala · Nachwuchssport · Talentidentifikation · Testkonstruktion · Psychologie

vated behavior? Furr and Funder (2010) state that in general personality psychology we know a lot about persons, but much less about situations and behaviors. In sport, to the best of our knowledge no theoretical approach exists that explicitly deals with achievement-motivated behavior. The concept that seems closest to us in this regard is self-regulation or volition. Volition is defined as “... meta-motivational processes required to sta-

bilize motivation in order to maintain the intended action, especially when rewards are not available immediately” (Elbe & Beckmann, 2006, p. 143). It is according to the Rubicon Model of Action Phases (at a glance: Achtziger & Gollwitzer, 2018) assuming that motivational processes only lead to the decision to act, and that volitional processes are then needed to implement the favored behavior (Achtziger & Gollwitzer, 2018).

Those processes are responsible for initiating an action and maintaining it until the goal has been reached, and include cognitive, motivational and emotional control strategies for not giving up when things get difficult, not letting oneself be distracted, not losing one's confidence, and staying positive (Elbe & Beckmann, 2006). As in the course of a sporting career, long-term goals must be pursued, and high training loads have to be completed, volition is of great importance for elite youth athletes (Baron-Thiene & Alfermann, 2015; Elbe, Szymanski, & Beckmann, 2005a). Volition or self-regulation can be assessed using self-rating questionnaires like the Volitional Components in Sport Questionnaire (Wenhold, Elbe, & Beckmann, 2008), which is a sport-specific adaptation of the general Volitional Components Questionnaire (Kuhl & Fuhrmann, 1998). The analysis of the items (e.g. "I am optimistic about most things in sports"; "In training, I often have to think about things that have nothing to do with what I'm doing") displays that also this construct, which should be closer located to behavior than the achievement motives, is not observable from the outside.

In summary, despite the very rich tradition of theories on achievement motives, it is not possible to deduce concrete performance-motivated behavior in sports from theory, which is why we have opted for an inductive approach for test construction (Smith, Fischer, & Fister, 2003).

Present research

To achieve our goal of constructing a reliable, valid and economical tool for the scientifically sound assessment of achievement-motivated behavior in sports, applicable in coaches every day practice (Horvath & Röthlin, 2018), we built upon a combination of a prototype (Broughton, 1984) and an inductive (Smith et al., 2003) test construction strategy. The act-frequency approach as a form of prototype approach relies on the definitions of constructs elaborated by psychological laypersons (Buss & Craik, 1983), and is used for item construction and to

provide evidence based on test content¹ (Studies 1 and 2). Despite the remarkable achievements in motivational psychology made in recent decades, little is known about observable achievement-motivated behaviors. To close this gap and to ensure that the tool is applicable within the practice of TID, in Study 1 (Act nomination) experienced coaches were asked about manifest achievement-motivated behavior in concrete situations (acts). In the subsequent phase (Study 2: Prototypicality rating), these acts were judged by other coaches and sport psychologists to assess evidence based on test content. In the third step (Study 3: Construction and initial validation), the final version of the rating scale was constructed and then checked for reliability and evidence based on internal structure and relation to a relevant criteria. The absence of an a priori theory on achievement-motivated behavior means that no factor or subscale structure is hypothesized in advance. Once the items were developed, the internal structure is uncovered, as when using inductive test construction strategy (Smith et al., 2003). Because it must be assumed that the achievement-motivated behavior of individual and team athletes differs due to their different settings (e.g. higher autonomy in individual sports, individual vs. team training), this study focuses on individual sports only. The ethics committee of the Phil.-hum. Faculty of the University of Bern approved this study.

Study 1: Act nomination

Method

Participants and procedure

The samples of the coaches were recruited directly through the sport federations via Swiss Olympic. The sporting directors of the sport federations categorized by Swiss Olympic in the categories 1 to 3 (of 5), according to their national importance

¹ We adopt the rather new standards for using the concept of validity and therefore do not use the historical nomenclatures like content or criterion validity (cf. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014)

and achievement potential, were asked to send lists of all their coaches at the 1st or 2nd level of education (professional training for elite sports or competitive sports). In total, we received the contact details of 438 coaches (15% female). For the studies in this publication, only German-speaking coaches were included. For act nomination, 36 coaches from 18 sport federations were randomly selected and invited to participate in Study 1. They received an informational letter with a link to the online survey. Overall, 20 coaches from 14 different sport federations ($M_{\text{age}} = 46.0 \pm 9.17$ years) took part and filled in the online survey. They had an average experience of $M = 16.15 \pm 9.63$ years as coaches in their sport.

Measures

In order to collect acts that the experienced coaches consider to represent achievement-motivated behavior, the following instructions were given in the online survey:

"Please think of one or more athletes whom you consider to be particularly achievement motivated with whom you are working with at the moment, or whom you have trained in the past. Now name three individual actions in specific situations in which the achievement motivation of these athletes you are thinking of was clearly expressed."

Data analysis

In this survey, the participants generated a total of 67 situations. These were minimally modified by the first author, without changing the content of the respective statements (cf. Amelang, Herboth, & Oefner, 1991; Krüger & Amelang, 1995): Very long and detailed situations were shortened (e.g. "He has carried out regenerative activities [black-roll or stretching] without being asked to do so" → "He/she has carried out regenerative activities without being asked to do so"), or divided into two acts (e.g. "He is very ambitious, sets himself goals and reacts emotionally to their achievement or failure" → "He/she has set goals" and "He/she has reacted emotionally when he/she failed to reach a self-imposed

goal”). Errors in grammar or spelling, as well as colloquial expressions were corrected. Repeatedly listed acts were deleted. All the modifications were extensively discussed with the coauthor to ensure that no unjustified modifications were made.

Results

Altogether, these adjustments resulted in 58 acts, which were included in the second study, where they were tested with regard to being prototypical of achievement-motivation behavior.

Study 2: Prototypicality rating

The goals of study 2 were twofold: On the one hand, evidence based on test content of the acts was assessed by a comparison between psychological laypersons (expert coaches) and experts in relation to achievement motivation (sport psychologists). On the other hand, by the examination of the rated prototypicality referring to achievement-motivated behavior of all acts, it should be avoided that especially low prototypical or inappropriate acts were processed further.

Method

Participants and procedure

For the prototypicality rating, 40 coaches from 18 sport federations and 150 sport psychologists were invited to participate in Study 2. The coaches were randomly selected and contacted via the contact lists mentioned earlier. The sport psychologists were contacted via the member list of the Swiss Association of Sport Psychology (SASP). They received an informational letter with a link to the online survey. In the end, 21 coaches from 12 different sport federations ($M_{\text{Age}} = 41.48 \pm 9.4$ years; $M_{\text{Experience}} = 14.20 \pm 5.67$ years), who had not participated in Study 1, and 26 sports psychologists ($M_{\text{Age}} = 43.23 \pm 10.14$ years; $M_{\text{Experience}} = 9.42 \pm 8.05$ years) took part.

Measures

The participants were asked to rate the 58 acts that had been constructed in

Study 1, based on the following instructions:

Below you will see a series of acts that may describe achievement motivation to a greater or a lesser extent. For each of these behaviors, please rate on a scale from 1 (not at all) to 5 (very) to what extent you consider this behavior to be achievement motivated, or whether in your opinion this act has something to do with achievement motivation.

In addition, the participants were asked to indicate their age, gender, vocational training and experience as a coach or sport psychologist.

Data analysis

The acts were presented in a random order. The prototypicality ratings were checked for mean differences between the two samples of the coaches and the sport psychologists. As the null hypothesis is the requested hypothesis, the alpha level was set to $\alpha = 0.20$. Cohen's d was calculated to determine the size of the effect.

Results

The results showed that the overall assessment leans towards *fairly prototypical* in both groups with $M = 3.75$ (standard deviation [SD] = 0.99). On the level of the individual acts, the range of prototypicality is between $M = 4.53$ (“He/she has set goals for himself/herself”) and $M = 2.74$ (“He/she cried when he/she did not win despite a good performance in the competition because he/she judged himself/herself better than the opponents”), with only five of 58 acts lying slightly below the level of 3 (*somewhat prototypical*). Out of the 58 acts, 14 showed significant differences between the two groups with $p < 0.20$. Out of this, the group judgments differ for nine acts with a large ($d > 0.8$) or moderate ($d > 0.5$) effect: Seven acts were regarded as more prototypical by the sports psychologists (e.g. “He/she has asked how he/she could further develop themselves in sports”). Only two acts were considered more prototypical by the coaches (e.g. “He/she strived to be the best at performance comparisons in another type of sport”). For all other

acts, there are no or only small group differences in terms of the prototypicality assessment.

Brief discussion of Study 2

The results show that, on average, the acts were rated as being fairly prototypical and therefore appropriate for measuring the concept of achievement-motivated behavior. For nine acts, the two groups do not agree about the prototypicality with a meaningful different assessment ($d > 0.5$). These items were removed from the rating scale in a first step because the results of the group comparison indicate that the evidence based on test content of these acts is unclear.² In terms of differences in the perception of the two samples, a trend became apparent that behavior pointing to the concept of task orientation (Duda, 2007), that is striving to achieve one's own goals and to constantly improve oneself, is regarded as more prototypical, and therefore of higher relevance for achievement motivated behavior by the sports psychologists than by the coaches. This indicates that the two groups might conceive the concept of achievement motivation in a slightly different form, or that the coaches make the everyday experiences that task orientation is not as important than it is considered to be in research (Elbe & Wikman, 2017).

As, to our knowledge, there exists no threshold for sufficient prototypicality, and as only five acts showed prototypicalities slightly lower than *somewhat prototypical*, all other acts were temporarily retained in the item pool (but none of the lower prototypical acts was included in the final rating scale). The upcoming validation phase had to examine whether the remaining 49 acts could be arranged to form a reliable and valid external rating scale for assessing achievement-motivated behavior.

² The other five acts which showed small differences were not removed in this step but all are excluded from the final instrument.

Study 3: Construction and validation

In Study 3, the first step (t_1) was to reduce the number of items, and to explore the underlying factor structure of the item pool to such an extent that at least three items per scale (Smith et al., 2003) create an economically applicable instrument for the measurement of the Achievement-Motivated Behavior in Individual Sports (AMBIS-I). The reduction of items seems for us to be a necessary precondition to ensure that the coaches, as the potential user of the coach-rating scale, will be willing to use the instrument in further practice (Horvath & Röthlin, 2018). To check the dimensionality seems constructive, as concerning the magnitude of different theories related to achievement motivation, a homogeneous 1-factor structure cannot be expected, and as the absence of an a priori theory within the inductive approach means that no factor structure is hypothesized in advance (Smith et al., 2003). The second step ($t_{2/3}$) checks whether the designed instrument fulfills the requirements towards reliability and internal structure. Finally, we checked the evidence for test-criterion relationship in terms of whether there are differences in the achievement-motivated behavior depending on the level of performance.

Method

Participants and procedure

The sample for Study 3 consisted of coaches who were asked to rate the achievement-motivated behavior of their athletes. Using the lists provided by the sport federations, 160 coaches were invited to participate in the investigation. 69 coaches participated in the study at t_1 , rating 288 athletes in individual sports with AMBIS-I. According to the guidelines of Swiss Olympic sports, disciplines comprising of teams with a small number of athletes (e.g. tennis, rowing) are counted in addition to individual sports. In a next step of the pre-analysis, the assessments of those coaches who indicated that they had known their athletes for less than half a year, or that they did not feel certain in their assess-

ment, were removed from the dataset. The data were also checked for extreme responding, where no evidence of occurrence was found. The final sample of the coaches for Study 3 at t_1 is therefore composed of 278 ratings by 67 coaches (19 women, 28.4%, 48 men, 71.6%, $M_{\text{age}} = 41.88 \pm 11.96$ years) from the following sports: badminton, biathlon, curling, free skiing, golf, judo, artistic cycling, cross country, track and field, mountain biking, road cycling, sledding, rowing, swimming, alpine skiing, shooting, tennis and vaulting. The coaches exhibit a high level of education, with more than 50% having successfully completed the highest or 2nd highest level. They have $M = 16.19 \pm 10.93$ years of professional experience and have known the athletes they judged for an average of $M = 4.11 \pm 3.45$ years. At t_2 (6–8 weeks after t_1) and t_3 (4 months after t_1), 46 coaches (16 women, 35%; 29 men, 65%; $M_{\text{age}} = 42.95 \pm 11.67$ years) participated and conducted ratings of 176 athletes for the second time. For 52 athletes, the assessments of two coaches (mostly head and assistant coaches) are available at the first measurement point to assess inter-rater reliability. All data were collected using an internet-based questionnaire (LimeSurvey Version 2.50). The coaches had the opportunity to pause and save their ratings to continue another day.

Measures

Achievement-motivated behavior. For Study 3, a coach-rating version was created from each of the 49 acts. A 4-point scale from 0 (*never*) to 3 (*often*) with a *not able to respond* option was used. For the second measurement time ($t_{2/3}$), the answer format was extended by the category 4 (*always*), as we found ceiling effects at t_1 (see [Table 2](#)). Because each coach had submitted a list with the names of athletes they were training at the 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 was inserted] display the acts mentioned below during the last 12 months?

Further variables. As mentioned before, 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 the rated athlete. Finally, the educational level of the coaches was assessed.

Performance level. As an external performance criterion to assess the evidence for test-criterion relationship, we checked at t_1 whether the athletes rated by their coaches on AMBIS-I were currently holding a Swiss Olympic Card (SOC). SOCs display achievements reached in competitions and can additionally be considered an expression of existing potential. The national sport associations award them according to their specific selection concept (0: *no SOC*; 1: *local SOC*; 2: *regional SOC*; 3: *national SOC*; 4: *international/elite SOC*). Out of this criterion, we composed two groups: One group consisting of athletes on a regional or lower level, and therefore with a lower attributed potential, and a second group consisting of athletes on a national or international level, and therefore with a higher attributed potential.

Data analysis

The analysis of the data was done using IBM SPSS Statistics 24 and Mplus Version 8 (Muthén & Muthén, 2017). In order to construct a tool that can be applied economically, a large part of the original 49 items had to be removed. For this exploratory procedure and to gather evidence based on internal structure, exploratory structural equation modeling (ESEM) was chosen. We aimed to construct a tool that unites at least three items per factor, features a good model fit and at the same time avoids content redundancy as much as possible. The statistical criteria for the choice and elimination of the items are based on Appleton, Ntoumanis, Quested, Viladrich, and Duda (2016) and Payne, Hudson, Akehurst, and Ntoumanis (2013).

Exploratory structural equation modeling (ESEM). The ESEM model for t_1 was estimated using the robust maximum likelihood (MLR) estimator method with

oblique geomin rotation, in line with the exploratory nature of the analysis (e.g. epsilon was fixed at 0.5). In doing so, missing values were estimated by means of the full information maximum likelihood (FIML) procedure. Overall, the proportion of missing data was 4.3%. Following the recommendations of Schermelleh-Engel, Moosbrugger, and Müller (2003), a good fit is indicated when $\chi^2/df \leq 2.00$; CFI (comparative fit index) and TLI (Tucker-Lewis Index) ≥ 0.97 ; and root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) ≤ 0.05 . To test the generalizability of the proposed model, the model was cross-validated with the data from t_2 and t_3 . Because the factor structure was already known at that time, target rotation was chosen as a more confirmatory approach (Marsh, Morin, Parker, & Kaur, 2014).

Reliability and validity. To estimate the reliability of the indicators, we computed squared multiple correlations (SMC). To estimate the reliability of the constructs, we calculated the composite reliability (CR) (Bagozzi & Yi, 2012) and the average variance extracted (AVE) (Fornell & Larcker, 1981). We used $SMC \geq 0.40$, $CR \geq 0.70$ and $AVE \geq 0.50$ as cut-offs for good reliabilities. To estimate discriminant evidence of the tool, the Fornell-Larcker criterion (Fornell & Larcker, 1981) was calculated. This criterion requires that the AVE of one factor should be higher than any squared correlation with another factor, assuming discriminant evidence at the factor level. To determine the test-retest reliability, reference is made to the data at t_2 . The coaches of this partial sample were invited to participate again after 8 weeks. Altogether, 64 estimates were made after an average time of 3.2 ± 0.72 months. The inter-rater reliability between head coach and assistant coach was determined separately for each training group (consisting of athletes assessed by the same two coaches) using the intra-class coefficient (ICC 2-way random effects model, absolute agreement), and then averaged over each training group. The interpretation is based on the recommendations of Koo and Li (2016), who classified ICCs lower

than 0.50 as poor, between 0.50 and 0.74 as sufficient, between 0.75 and 0.89 as good and higher than 0.90 as excellent. To compare the extent of achievement-motivated behavior between the two performance groups, an independent t-test was performed using the values of the factors of AMBIS-I as dependent variables. The level of significance was set at $p < 0.05$ for all analyses.

Results

In a first step, we removed 24 items which were answered by 12 to 42% of the coaches with *not able to respond*. It was assumed that if a proportion of more than 10% of all the coaches stated that they could not answer, this is not an incidental finding, and that the reasons for not answering make such items not suitable in the rating scale. The following reasons are conceivable: (a) not directly observable (e.g. “He/she went home after the training to continue practicing there”, 35% did not answer), (b) unsuitable for certain sports (e.g. “He/she cried after a final defeat”, 22%), or (c) improper evaluation period (e.g. “He/she decided to attend a specialized sports school, even though his/her friends continued to attend the regular one”; 15%; “He/she has moved his/her place of residence closer to the training location”, 20%). As the items were all constructed by coaches without any experience in item and test construction (study 1), those problematic points were not addressed automatically, as would have been addressed if the researchers themselves would have constructed the items.

Construction of AMBIS-I

For the remaining 25 items, six ESEM models with 1 to 6 factors were modeled and tested. The comparison of the different models points to the 3 or 4 factor model, as these two models have the lowest BIC (Schwarz’s Bayesian Information Criteria), and thus display a better model fit taking into account the principle of economy (Masyn, 2013). In both the 3- and 4-factor versions, incremental items were then removed if they featured main loadings ≤ 0.50 , communalities ≤ 0.40 , high cross loadings or con-

tent redundancy. After multiple phases of ESEM modeling and item removal, a 3-factor model with 12 items displaying an excellent fit (CFI = 0.99, TLI = 0.97 and RMSEA = 0.04) and readily interpretable factors emerged as the best model (see [Table 1](#), AMBIS-I-12 t_1).

For the cross-validation, the data from t_2 and t_3 was used to check the proposed 12-item solution. The AMBIS-I-12 $t_{2/3}$ solution displays an inferior fit with the data (cf. AMBIS-I-12 $t_{2/3}$, [Table 1](#)). Because act 03 showed a low loading on its factors and therefore small communality, and as act 47 showed high cross-loadings, we decided to remove these two items, resulting in an excellent match between the estimated and the observed data at t_1 (cf. AMBIS-I-10 t_1 , [Table 1](#)), as well as at $t_{2/3}$. The coach-rating scale for achievement-motivated behavior in individual sports (AMBIS-I) therefore consists of 10 items forming three factors which were according to their content (see discussion), named proactivity, ambition and commitment.

As shown by the descriptive statistics in [Table 2](#), proactivity is least likely to be displayed, followed by ambition and commitment. Additionally, commitment exhibits restricted variance, especially at t_1 , and thus probably a ceiling effect. The higher values in all dimensions at $t_{2/3}$ compared with t_1 are at least partly due to the change in the scaling from 0–3 at t_1 to 0–4 at $t_{2/3}$. The factor intercorrelations lie in the range $0.22 < r < 0.67$, which is acceptable with regard to the differentiation of content ([Table 2](#)). Due to the significant correlations of the three factors, especially at $t_{2/3}$, the factors cannot be assumed to be uncorrelated. For this reason, the reliability of a total score—in the sense of a higher-order factor—will be examined as well. The disadvantage of the lower content specificity of a total score could possibly be compensated for by improved reliability (through the higher length of the test) and thereby associated with higher validity (Furr & Bacharach, 2014; Kingston, Scheuring, & Kramer, 2013).

Table 1 Goodness of fit statistics and information criteria for the models estimated in Study 3

Models	χ^2	p (df)	χ^2/df	CFI	TLI	RMSEA (CI 90%)	SRMR
Accepted standard	–	>0.05	≤2	<0.97	<0.97	≤0.05	≤0.05
AMBIS-I-12 t_1	44.20	0.09 (33)	1.34	0.99	0.97	0.04 (0.00–0.06)	0.02
AMBIS-I-12 $t_{2/3}$	64.64	<0.001 (33)	1.96	0.96	0.92	0.07 (0.05–0.10)	0.03
AMIBIS-I-10 t_1	18.94	0.40 (18)	1.05	1.00	1.00	0.01 (0.00–0.06)	0.02
AMBIS-I-10 $t_{2/3}$	26.09	0.10 (47)	1.45	0.99	0.97	0.05 (0.00–0.09)	0.05

CFI comparative fit index, TLI Tucker-Lewis index, RMSEA root mean square error of approximation, SRMR standardized root mean square residual

Table 2 Descriptive statistics and intercorrelations of the AMBIS-I factors and the total score at t_1 and $t_{2/3}$

Factors		Descriptive statistics		Intercorrelations		
		<i>M</i>	<i>SD</i>	Proactivity	Ambition	Commitment
Proactivity	t_1	1.51	0.75	1	–	–
	$t_{2/3}$	1.78	0.93	1	–	–
Ambition	t_1	1.80	0.82	0.22	1	–
	$t_{2/3}$	2.24	0.95	0.55	1	–
Commitment	t_1	2.40	0.53	0.30	0.23	1
	$t_{2/3}$	2.81	0.68	0.67	0.39	1
Total score AMBIS-I	t_1	1.91	0.53	–	–	–
	$t_{2/3}$	2.28	0.71	–	–	–

t_1 : Scale 0–3; $t_{2/3}$: Scale 0–4, *M* mean, *SD* standard deviation

Reliability and evidence based on internal structure

The corresponding items of AMBIS-I, their factor loadings and communalities, and the factor reliabilities for t_1 and $t_{2/3}$ are displayed in Table 3. It should be noted that the items were constructed and checked in German, and that the English translations of the items in Table 3 have not been validated and are only included for informational purposes. The original German wording of the items is available in Table 1 of the Electronic Supplementary Material 1. The factor loadings on the main factors lie between 0.48 and 0.84 for all items. Five items exhibit cross loadings > 0.20. C1 has a somewhat low indicator reliability at t_1 . All other items display good communalities (>0.40). The two factors proactivity and ambition display good factor reliabilities at both times of measurement. Only commitment has factor reliabilities that lie below the desired thresholds, especially at t_1 . In the cross validation at $t_{2/3}$, their CR = 0.68 and AVE = 0.49 values lie only just below the desired threshold values of 0.70 and 0.50. The test–retest reliability r_{tt} is

≥0.70 for all three factors and thus within an acceptable range over a comparatively long period of 3.2 ± 0.72 months. The total score has the highest retest reliability, with $r_{tt} = 0.79$. As the average variance extracted for each factor is higher than any squared correlation with another factor, all three factors meet the Fornell-Larcker criterion and therefore display discriminant evidence at the factor level.

Inter-rater reliability

The intra-class coefficients (ICCs) of the coaches averaged over all training groups are displayed in Table 4. All ICCs except that for proactivity lie in a satisfactory range ($0.51 < ICC < 0.70$). Generally, the average measure, i.e. ICCs which use the mean values of the two coaches as a source of information, have higher values. The inter-rater reliabilities of the total score show the highest ICCs for both the average and the single rater measures.

Evidence for test–criterion relationship

In matters of concurrent evidence for test–criterion relationship, all three factors, as

well as the total score, differentiate with small to medium effects ($0.30 < d < 0.62$) between the two performance groups determined by the SOC (Table 4). Athletes on the national or international level were rated as showing more achievement-motivated behavior than those on the lower performance level.

Brief discussion of Study 3

In the first step of Study 3, we simultaneously reduced the number of items to ten and checked the dimensionality. The final 10-item tool AMBIS-I, mapping the three factors proactivity, ambition and commitment, showed excellent fit indices and was successfully replicated by the data at $t_{2/3}$.

The factor *proactivity* refers to getting involved in training processes on one's own initiative and for one's own sake. The impulse to act comes from the person itself, and an action is carried out self-determined and does not require an external push (e.g. item P4: The athlete looks for opportunities to catch up on missed training content). Self-determined means that motivation emanates from the self (Ryan & Deci, 2000). Intrinsic motivation as the most self-determined motivation, and thus accompanied with proactivity, is characterized by the sensation of immanent pleasure stemming from performing an activity itself. Another theoretical link can be made to the concept of achievement goal orientations: A proactive athlete pursues his or her goals perseveringly (e.g. item P2: Athletes to be first on the training grounds and practicing technical processes independently). The items of the factor proactivity give no indication of which goals are being closely pursued. The fact that those goals are being pursued persistently furthermore makes the connection with volition obvious.

The factor *ambition* is characterized by the absolute will to successfully pursue self-imposed goals in competitions. Ambitious athletes aim at winning competitions (e.g. item A2: The athlete clearly communicates before the competition that he/she wanted to win). If those goals were not achieved, the athlete responds unsatisfied (e.g. item A1:

Table 3 Factor loadings, subscales and communalities of items, and factor reliability of subscales and retest reliabilities of subscales and total score

Scales/Items	F1		F2		F3		SMC		CR		AVE		R _{tt}
	t ₁	t _{2/3}											
Total Score AMBIS-I													0.79
1. Proactivity									0.79	0.86	0.54	0.63	0.74
P1: He/she urged the coach to plan further training sessions to improve even further	0.64	0.83	0.21	–	–	–	0.48	0.62	–	–	–	–	–
P2: He/she was the first to be on the training grounds and practice technical processes independently	0.74	0.76	–	–	–	–	0.64	0.67	–	–	–	–	–
P3: He/she stayed on after training to continue practicing	0.73	0.78	–	–	–	–	0.60	0.65	–	–	–	–	–
P4: He/she independently looked for opportunities to catch up on missed training content	0.50	0.71	–	–	0.32	–	0.46	0.57	–	–	–	–	–
2. Ambition									0.79	0.87	0.59	0.71	0.71
A1: He/she has acted annoyed when he/she did not finish the competition in first place	–	–	0.68	0.82	–	–	0.51	0.64	–	–	–	–	–
A2: He/she clearly communicated before the competition that he/she wanted to win today	–	–	0.71	0.84	0.21	–	0.59	0.80	–	–	–	–	–
A3: He/she showed that he/she is not satisfied with coming in 2nd place	–	–	0.77	0.78	–	–	0.66	0.68	–	–	–	–	–
3. Commitment									0.62	0.68	0.40	0.49	0.70
C1: He/she oriented himself/herself towards stronger athletes during training	–	–	–	0.24	0.70	0.48	0.38	0.53	–	–	–	–	–
C2: In highly demanding exercises, he/she worked to the point of exhaustion	–	–	–	–	0.62	0.57	0.41	0.41	–	–	–	–	–
C3: He/she has shown an “active” stance in training	–	–	0.26	–	0.59	0.75	0.42	0.51	–	–	–	–	–

F1–F3 Factor loadings; loadings < 0.20 are not displayed, SMC estimated squared multiple correlation or communality of the item, CR Composite reliability; AVE average variance extracted. The original German items are available in Table 1 of Electronic Supplementary Material 1

The athlete has acted annoyed when he did not finish the competition in first place.) Thus, the conceptual proximity of the factor ambition to achievement goal orientations competitiveness and goal orientation (as the goals are self-imposed), and—to a lesser degree—ego orientation, becomes visible.

In turn, the factor *commitment* is again localized in the setting of training. Committed athletes show that they are ready and willing to perform (e.g., item C3: The athlete shows an “active” stance in training). He or she displays will to work hard to achieve a goal and tries to solve a task repeatedly, even in the face of adversity (Scanlan, Carpenter, Simons, Schmidt, & Keeler, 1993). Hence, commitment is closely related to discipline

in training (e.g. item C1: In highly demanding exercises, the athlete works to the point of exhaustion) and volition, as volition is especially important for “realizing long and intense training loads during the course of an athletic career or for keeping up regular exercising” (Elbe et al., 2005a, p. 560). Because committed athletes orient towards and compare themselves to stronger athletes, a connection with competitiveness is also to be assumed (Gill & Deeter, 1988).

In the second step, we checked whether the designed instrument fulfills initial requirements towards its psychometric properties. Overall, the results point to satisfactory reliability and evidence based on internal structure. The factor reliability of commitment is a bit

more critical. However, as commitment shows the highest width in terms of content, and at the same time the lowest variance (e.g. Table 2), this is hardly surprising. Contrasting these results, commitment showed the best values for inter-rater reliability. Although the values for the inter-rater reliability are not very high in the overview, the results coincide with coefficients of a meta-analysis of the inter-rater reliability of supervisor ratings in an occupational context (Conway & Huffcutt, 1997). Besides, to increase the reliability of the statements, the coach assessment should where possible be carried out by two coaches (Musculus & Lobinger, 2018) and the values should then be averaged for further interpretation. The findings

Table 4 Intraclass correlation coefficients (ICC) as measure of agreement between head and assistant coach averaged over all training groups ($n = 13$) for t_1 (ICC 2-way random effects model, absolute agreement) and descriptive statistics and between-group analyses of AMBIS-I

Scales	Intraclass correlation		Regional level & lower ($n = 159$)		National level & higher ($n = 119$)		t (df)	p	d (95% CI)
	ICC (single rater)	ICC (mean of 2 raters)	M	SD	M	SD			
Proactivity	0.42	0.55	1.32	0.77	1.77	0.65	5.23 (270)	<0.001	0.62 (0.38–0.87)
Ambition	0.51	0.65	1.69	0.83	1.93	0.78	2.44 (268)	0.02	0.30 (0.06–0.54)
Commitment	0.58	0.70	2.33	0.55	2.49	0.49	2.48 (276)	0.01	0.31 (0.07–0.54)
Total score AMBIS-I	0.60	0.72	1.78	0.55	2.07	0.46	4.65 (276)	<0.001	0.57 (0.32–0.81)

The 13 training groups (athletes and two coaches) comprise two to eight athletes. Nine training groups even comprise less than five athletes
 M mean, SD standard deviation, CI confidence interval

on test-criterion relationship point in the expected direction, and are—in the case of ambition and commitment—in line with previous findings from a self-report study on the difference between elite- and subelite-level athletes (Halldorsson, Helgason, & Thorlindsson, 2012). The magnitude of the uncovered differences—with small to medium effects—is substantial considering performance in sports is determined in a multidimensional manner. Thus, in addition to psychological characteristics, such as achievement-motivated behavior, features from various other areas are also relevant, including motor performance, social support or the characteristics of training (Meyer, Gnacinski, & Flechter, 2018; Rees et al., 2016).

General discussion

The studies presented aimed to construct a reliable, valid and economic tool for the scientifically sound assessment of achievement-motivated behavior in sports. This stemmed from the fact that the tools currently available for assessing achievement motivation display methodological problems such as social desirability of self-rating questionnaires, motives that were not directly accessible by coaches, and high costs if implicit motives should be assessed. Those three limitations in the field of selection decisions in youth elite sports can be reduced by using AMBIS-I. With AMBIS-I, socially desirable responding is less problematic as the assessment

is based on coach-ratings instead of the ratings of the athletes themselves, and the focus is switched from motives to observable behaviors. The items generated in Study 1 refer to specific situations, in which the coaches' lay understanding of achievement motivation was clearly expressed in the behavior of their athletes.

In Study 2, these items were judged by sport psychologists as experts in regard to achievement motivation, and by coaches as psychological laypersons in terms of their being content-valid/prototypical for the construct of achievement-motivated behavior. The results of Study 3 then indicate that AMBIS-I is an economic tool that meets many requirements of reliability and validity. The results can be summarized as follows: The external rating scale AMBIS-I contains the three factors proactivity, ambition and commitment based on ten forms of behavior, whose frequency of occurrence is rated by the coaches. The three factors display an excellent fit with the observed data, reasonable to good reliability coefficients and good evidence based on internal structure. The first findings on test-criterion relationship are promising when considering the potential use of AMBIS-I as an instrument in a battery of tests for talent selection in the future.

Due to the larger number of items, the total score is somewhat more reliable than the individual factors (Tavakol & Dennick, 2011). The calculation of a total score also makes sense from the point of view of content, as in Study 1 the coaches

were questioned about achievement-motivated behavior in general and not about forms of behavior representing the retrospectively extracted factors proactivity, ambition or commitment. We therefore recommend using the individual factors as well as the overall score for interpretation. The individual factors provide more detailed information, whereas, from a statistical point of view, the overall score offers more reliable statements. AMBIS-I can therefore be described as a multidimensional test with correlated factors (according to the definition of Furr & Bacharach, 2014).

Because future users were included in the construction of the items and in assessing the quality of the items, it can be assumed that the final tool will be well accepted in practice. We recommend using AMBIS-I as part of a multidimensional test battery for the TID. From the point of view of talent development and further career planning, it might also be interesting to compare the coaches' ratings with the self-ratings of the athletes, and discuss possible differences or similarities (Musculus & Lobinger, 2018).

Despite the encouraging results so far, some critical points need to be addressed. The cross-validation with the same sample at a later date is not the gold standard; however, due to the limited sample size and accessibility of the coaches responsible for selection decisions, this seems to be an acceptable solution. Furthermore, the need to adapt the scaling of the AMBIS-I items between the first and the second measurement time is not ideal.

Nevertheless, the successful cross-validation and the improved distribution of the values at $t_{2/3}$ confirm the decision to implement this adaptation. As it is also not possible for other test constructions, we cannot guarantee that the full breadth of achievement-motivated behavior is covered. However, there were a large number of acts nominated in Study 1 and many of them were also overlapping each other which suggests that the most important situations might be mentioned. Additionally, Smith et al. (2003, p. 473) argue that inductive test construction has the advantage that "... investigators are likely to be able to identify all or most elements of the construct that are recognized by the target individuals. The approach lessens the possibility that a limited theoretical perspective will cause investigators to omit important construct facets".

In contrast to motor tests, there are no purely objective measured values when recording psychological characteristics. Although the social desirability of the self-ratings can be reduced with the coach-ratings used, not all possible answering tendencies (e.g. recall or confirmation bias; Althubaiti, 2016) can be controlled for even though focusing on specific behavior might not be as strongly affected by some response biases as other forms of assessment (McCrae & Weiss, 2010). The fact that the number of athletes to be rated by one coach is low speaks against a large effect of systematic errors caused by differences in individual standards of the raters. The same is true concerning our results for inter-rater reliability, which lie in an acceptable range and therefore speak against a high impact of individual effects of—for example—confirmation bias. To avoid primacy or recency effects, the order of athletes to be rated was randomly changed between the first and the second measurement point, which resulted in acceptable retest reliabilities. In addition, we should consider the whole talent selection process, for which it is strongly recommended to use multidimensional test batteries (motor performance, social support or the characteristics of training), and therefore also different data collection instruments as motor tests or performance results (Vaeyens et al.,

2008). Hence, selection decisions should not be based on just one instrument or variable. Therefore, the effect of possible rating biases from AMBIS-I on the holistic selection decision can thus be considered small. However, whenever possible, it is recommended to have the ratings of two coaches (cf. main and assistant coach) to mitigate subjective preferences and confirmation bias.

Despite the promising results based on content, internal structure and test-criterion relationship, the next step has to be the ongoing validation of the test. On the one hand, relationships between test scores and other measures must be carried out by checking whether AMBIS-I is related to other instruments measuring motivational and volitional constructs and therefore refer to the underlying construct achievement motives. From a content perspective, connections can be assumed to exist between proactivity and intrinsic motivation (Deci & Ryan, 1985), or between ambition and competitiveness (Gill & Deeter, 1988). In addition, it seems reasonable to examine the extent to which self-assessment regarding achievement-motivated behavior is related to the coaches' assessment. Although high correlations are not be expected (Conway & Huffcutt, 1997), it seems important to review the overall trend and to involve different perspectives. On the other hand, there are still important questions regarding test-criterion relationship. Thus, an instrument that is meant to predict athletic performance or success should not only be examined using a cross-sectional design, as was done in the present study, but also in a prognostic manner (Vaeyens et al., 2008). Both coach assessment and objective success criteria can be used as performance criteria.

In addition to the extended validation, it also makes sense to check the retest reliability over a shorter time period. Participants took an average of more than 3 months to complete the online questionnaire, despite an early reminder for participation at the second date of measurement. Therefore the calculated reliability coefficient reflects, on the one hand, the degree to which measurements error affects the test and, on the other hand,

the amount of change in the true scores (Furr & Bacharach, 2014). It must therefore be regarded as a mixture between the examination of stability and reliability. In the literature, many test-retest analyses are conducted over a period of 2 to 8 weeks (Furr & Bacharach, 2014).

Besides curiosity about learning more about the psychometric properties, advanced analyses of stability, change and age-relatedness of achievement-motivated behavior are of great interest. It can, for example, be assumed that very high values in proactivity are hard to achieve for the younger athletes, as they might—due to external circumstances—be unable to decide for themselves when to arrive at training, and will therefore be accompanied by parents. With the progressive development of behavioral autonomy, autonomous actions and decisions, higher values will be possible as the athletes move toward young adulthood (Collins, Gleason, & Sesma, 1997). The question of age-dependent standard values should therefore be addressed in the near future. In addition, it is necessary to examine whether the much-expressed assumption of the highest possible values in achievement-motivated behavior must be rated positively in every case, or whether extreme manifestations, for example in proactivity, can be accompanied by negative phenomena, such as over-involvement or burnout (Gardner & Moore, 2006). Additionally, it should be emphasized once again that AMBIS-I has so far only been tested for individual sports, and that there has been no validation yet for team sports. So, it is essential for future studies to check if the items of AMBIS-I are also suitable for team sports.

Based on the current state of knowledge, the newly constructed, economically usable coach-rating scale AMBIS-I allows a reliable assessment of observable achievement-motivated behavior in individual sports. Additionally, promising initial indications of validity can be reported. On this basis, a first step has been taken in identifying future successful athletes early on and accordingly promoting them more effectively.

Corresponding address



Claudia Zuber
Universität Bern
Bern, Switzerland
claudia.zuber@ispw.unibe.ch

Compliance with ethical guidelines

Conflict of interest C. Zuber and A. Conzelmann declare that they have no competing interests.

All procedures performed in studies involving human participants or on human tissue were in accordance with the ethical standards of the institutional and/or national research committee and with the 1975 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abbott, A., & Collins, D. (2004). Eliminating the dichotomy between theory and practice in talent identification and development: considering the role of psychology. *Journal of Sports Sciences*, 22(5), 395–408.
- Achtziger, A., & Gollwitzer, P. M. (2018). Motivation and volition in the course of action. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation and Action* (3rd edn., pp. 485–527). Cham: Springer.
- Althubaiti, A. (2016). Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of Multidisciplinary Healthcare*, 9, 211–217. <https://doi.org/10.2147/JMDH.S104807>.
- Amelang, M., Herboth, G., & Oefner, I. (1991). A prototype strategy for the construction of a creativity scale. *European Journal of Personality*, 5, 261–285. <https://doi.org/10.1002/per.2410050402>.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: students' learning strategies and motivation processes. *Journal of Educational Psychology*, 80, 260–267. <https://doi.org/10.1037/0022-0663.80.3.260>.
- Anshel, M., & Lidor, R. (2009). Talent detection programs in sport: The questionable use of psychological measures. *Journal of Sport Behavior*, 35(3), 239–266.
- Appleton, P. R., Ntoumanis, N., Quedest, E., Viladrich, C., & Duda, J. L. (2016). Initial validation of the coach-created empowering and disempowering motivational climate questionnaire (EDMCQ-C). *Psychology of Sport and Exercise*, 22, 53–65.
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, 64(6), 359–372.
- Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science*, 40(1), 8–34.
- Baron-Thiene, A., & Alfermann, D. (2015). Personal characteristics as predictors for dual career dropout versus continuation—A prospective study of adolescent athletes from German elite sport schools. *Psychology of Sport and Exercise*, 21, 42–49.
- Broughton, R. (1984). A prototype strategy for construction of personality scales. *Journal of Personality and Social Psychology*, 47(6), 1334–1346. <https://doi.org/10.1037/0022-3514.47.6.1334>.
- Brunstein, J. C., & Heckhausen, H. (2010). Achievement motivation. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation and action* (1st edn., pp. 137–183). Cambridge: University Press.
- Burtscher, J., Furtner, M., Sachse, P., & Burtscher, M. (2011). Validation of a German version of the Sport Motivation Scale (SMS28) and motivational analysis in competitive mountain runners. *Perceptual and motor skills*, 112(3), 807–820. <https://doi.org/10.2466/05.06.25.PMS.112.3.807-820>.
- Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review*, 90(2), 105–126.
- Calvo, T. G., Cervelló, E., Jiménez, R., Iglesias, D., & Murcia, J. A. M. (2010). Using self-determination theory to explain sport persistence and dropout in adolescent athletes. *Spanish Journal of Psychology*, 13, 677–684.
- Christensen, M. K. (2009). “An eye for talent”: talent identification and the “practical sense” of top-level soccer coaches. *Sociology of Sport Journal*, 26(3), 365–382.
- Coetzee, B., Grobbelaar, H. W., & Gird, C. C. (2006). Sport psychological skills that distinguish successful from less successful soccer teams. *Journal of Human Movement Studies*, 51(6), 383–401.
- Collins, W., Gleason, T., & Sesma, A. (1997). Internalization, autonomy, and relationships: Development during adolescence. In J. E. Grusec & L. Kuczynski (Eds.), *Parenting and children's internalization of values. A handbook of contemporary theory* (pp. 78–99). New York: John Wiley & Sons.
- Connelly, B. S., & Ones, D. S. (2010). Another perspective on personality: meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, 136(6), 1092–1122. <https://doi.org/10.1037/a0021212>.
- Conway, J. M., & Huffcutt, A. I. (1997). Psychometric properties of multisource performance ratings: a meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance*, 10(4), 331–360. https://doi.org/10.1207/s15327043hup1004_2.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Duda, J. (2007). Motivation in sport: the relevance of competence and achievement goals. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 318–335). New York: Guilford.
- Elbe, A.-M. (2004). Testgütekriterien der deutschen Version des Sport Orientation Questionnaires [Psychometric properties of the German version of the Sport Orientation Questionnaire]. *Spectrum der Sportwissenschaft*, 16(1), 96–107.
- Elbe, A.-M., & Beckmann, J. (2006). Essential processes for attaining peak performance. In D. Hackfort & G. Tenenbaum (Eds.), *Motivational and self-regulatory factors and sport performance in young elite athletes* (pp. 137–157). Aachen: Meyer & Meyer Sport.
- Elbe, A.-M., & Wikman, J. M. (2017). Psychological factors in developing high performance athletes. In J. Baker, S. Copley, J. Schorer & N. Wattie (Eds.), *Routledge handbook of talent identification and development in sport* (pp. 169–178).
- Elbe, A.-M., Szymanski, B., & Beckmann, J. (2005a). The development of volition in young elite athletes. *Psychology of Sport and Exercise*, 6, 559–569. <https://doi.org/10.1016/j.psychsport.2004.07.004>.
- Elbe, A.-M., Wenhold, F., & Müller, D. (2005b). Zur Reliabilität und Validität der Achievement Motives Scale-Sport [On the reliability and validity of the Achievement motives scale—Sport]. *Zeitschrift für Sportpsychologie*, 12(2), 57–68.
- Elliot, A. J., & Dweck, C. S. (2007). Competence and motivation: Competence as the core of achievement motivation. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 3–12). New York: Guilford.
- Feichtinger, P., & Höner, O. (2014). Psychological diagnostics in the talent development program of the German Football Association: Psychometric properties of an Internet-based test battery. *Sportwissenschaft*, 44(4), 203–213. <https://doi.org/10.1007/s12662-014-0341-0>.
- Figueiredo, A. J., Gonçalves, C. E., Coelho, S. M., & Malina, R. M. (2009). Characteristics of youth soccer players who drop out, persist or move up. *Journal of Sports Sciences*, 27, 883–891.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39. <https://doi.org/10.2307/3151312>.
- Fuchslocher, J., Romann, M., Rüdüsüli, R., Birrer, D., & Hollenstein, C. (2011). Das Talentidentifikationsinstrument PISTE—Wie die Schweiz Nachwuchssportler auswählt [The talent identification tool PISTE—how Switzerland identifies junior athletes]. *Leistungssport*, 4(2), 22–27.
- Furr, R. M., & Bacharach, V. R. (2014). *Psychometrics: an introduction*. Los Angeles: SAGE.
- Furr, R. M., & Funder, D. C. (2010). Behavioral observation. In R. W. Robins, R. C. Fraley & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 273–291). New York, London: Guilford.
- Gardner, F., & Moore, Z. (2006). *Clinical sport psychology*. Champaign: Human Kinetics.
- Gill, D. L., & Deeter, T. E. (1988). Development of the sport orientation questionnaire. *Research Quarterly for Exercise and Sport*, 59(3), 191–202.
- Gillet, N., Berjot, S., Vallerand, R., Amoura, S., & Rosnet, E. (2012). Examining the motivation-performance relationship in competitive sport: a cluster-analytic approach. *International Journal of Sport Psychology*, 43, 79–102.
- Gillet, N., Vallerand, R., & Paty, B. (2013). Situational motivational profiles and performance with elite

- performers. *Journal of Applied Social Psychology*, 43, 1200–1210. <https://doi.org/10.1111/jasp.12083>.
- Gillet, N., Vallerand, R., & Rosnet, E. (2009). Motivational clusters and performance in a real-life setting. *Motivation and Emotion*, 33, 49–62.
- Halldórsson, V., Helgason, A., & Thorlindsson, T. (2012). Attitudes, commitment and motivation amongst Icelandic elite athletes. *International Journal of Sport Psychology*, 43, 241–254.
- Halvari, H., & Thomassen, T.O. (1997). Achievement motivation, sports-related future orientation, and sporting career. *Genetic Social and General Psychology Monographs*, 123, 343–365.
- Heckhausen, H., & Heckhausen, J. (2010). Motivation and action: Introduction and overview. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation and action* (pp. 1–9). Cambridge: University Press.
- Hellandsg, E. (1998). Motivation predictors of high performance and discontinuation in different types of sports among talented teenage athletes. *International Journal of Sport Psychology*, 29, 27–44.
- Höner, O., & Feichtinger, P. (2016). Psychological talent predictors in early adolescence and their empirical relationship with current and future performance in soccer. *Psychology of Sport and Exercise*, 25, 17–26.
- Horvath, S., & Röthlin, P. (2018). How to improve athletes' return of investment: Shortening questionnaires in the applied sport psychology setting. *Journal of Applied Sport Psychology*, 30(2), 241–248.
- Huijgen, B. C. H., Elferink-Gemser, M. T., Lemmink, K. A. P. M., & Visscher, C. (2014). Multidimensional performance characteristics in selected and deselected talented soccer players. *European Journal of Sport Science*, 14, 2–10. <https://doi.org/10.1080/17461391.2012.725102>.
- Johnson, J. A. (1997). Units of analysis for the description and explanation of personality. In R. Hogan, J. Johnson & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 73–93). San Diego: Academic Press.
- Johnston, K., Wattie, N., Schorer, J., & Baker, J. (2018). Talent identification in sport: a systematic review. *Sports Medicine*, 48(1), 97–109. <https://doi.org/10.1007/s40279-017-0803-2>.
- Jokuschies, N., Gut, V., & Conzelmann, A. (2017). Systematizing coaches' 'eye for talent': player assessments based on expert coaches' subjective talent criteria in top-level youth soccer. *International Journal of Sports Science & Coaching*, 12(5), 565–576. <https://doi.org/10.1177/1747954117727646>.
- Kingston, N. M., Scheuring, S. T., & Kramer, L. B. (2013). Test development strategies. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise & M. C. Rodriguez (Eds.), *Test theory and testing and assessment in industrial and organizational psychology*. APA handbook of testing and assessment in psychology, (Vol. 1, pp. 165–184). Washington: American Psychological Association.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163.
- Krüger, C., & Amelang, M. (1995). Bereitschaft zu riskantem Verhalten als Trait-Konstrukt und Test-Konzept: Zur Entwicklung eines Fragebogens auf der Basis des Handlungs-Häufigkeits-Ansatzes [Willingness for risk-taking behavior as trait construct and test concept: The construction of a questionnaire based on the act-frequency approach]. *Diagnostica*, 41, 35–52.
- Kuhl, J., & Fuhrmann, A. (1998). Decomposing self-regulation and self-control: The volitional components inventory. In J. Heckhausen & C. S. Dweck (Eds.), *Motivation and self-regulation across the life span* (pp. 15–99). Cambridge: Cambridge University Press.
- MacNamara, Á., Button, A., & Collins, D. (2010). The role of psychological characteristics in facilitating the pathway to elite performance: Part 1: Identifying mental skills and behaviors. *The Sport Psychologist*, 24, 52–73.
- Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, 10, 85–110.
- Masyn, K. (2013). Latent class analysis and finite mixture modeling. In T. D. Little (Ed.), *Oxford library of psychology. The Oxford handbook of quantitative methods in psychology* (pp. 551–611). New York: Oxford University Press.
- McCrae, R. R., & Weiss, A. (2010). Observer ratings of personality. In R. W. Robins, R. C. Fraley & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 259–272). New York, London: Guilford.
- Meyer, B. B., Gnacinski, S. L., & Flechter, T. B. (2018). Talent identification. In J. Taylor (Ed.), *Assessment in applied sport psychology* (pp. 169–182). Champaign: Human Kinetics.
- Murr, D., Feichtinger, P., Larkin, P., O'Connor, D., & Höner, O. (2018). Psychological talent predictors in youth soccer: a systematic review of the prognostic relevance of psychomotor, perceptual-cognitive and personality-related factors. *PLOS ONE*, 13, e205337.
- Musculus, L., & Lobinger, B. H. (2018). Psychological characteristics in talented soccer players—Recommendations on how to improve coaches' assessment. *Frontiers in Psychology*, 9, 395.
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide*. Los Angeles: Muthén & Muthén.
- Nicholls, J. G. (1984). Achievement motivation: conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91(3), 328–346. <https://doi.org/10.1037/0033-295X.91.3.328>.
- Payne, S. M., Hudson, J., Akehurst, S., & Ntoumanis, N. (2013). Development and initial validation of the impression motivation in sport questionnaire-team: team. *Journal of Sport and Exercise Psychology*, 35(3), 281–298. <https://doi.org/10.1123/jsep.35.3.281>.
- Pelletier, L. G., Fortier, M. S., Vallerand, R. J., & Brière, N. M. (2001). Associations among perceived autonomy support, forms of self-regulation, and persistence: a prospective study. *Motivation and Emotion*, 25, 279–306.
- Pelletier, L. G., Rocchi, M. A., Vallerand, R. J., Deci, E. L., & Ryan, R. M. (2013). Validation of the revised sport motivation scale (SMS-II). *Psychology of Sport and Exercise*, 14(3), 329–341. <https://doi.org/10.1016/j.psychsport.2012.12.002>.
- Pelletier, L. G., Tuson, K. M., Fortier, M. S., Vallerand, R. J., Brière, N. M., & Blais, M. R. (1995). Toward a new measure of intrinsic motivation, extrinsic motivation, and amotivation in sports: The Sport Motivation Scale (SMS). *Journal of Sport and Exercise Psychology*, 17(1), 35–53.
- Rees, T., Hardy, L., Güllich, A., Abernethy, B., Côté, J., Woodman, T., Warr, C., et al. (2016). The Great British Medalists Project: a review of current knowledge on the development of the world's best sporting talent. *Sports Medicine*, 46(8), 1041–1058. <https://doi.org/10.1007/s40279-016-0476-2>.
- Reilly, T., Williams, M., Nevill, A., & Franks, A. M. (2000). A multidisciplinary approach to talent identification in soccer. *Journal of Sports Sciences*, 18, 695–702.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, 55, 68–78.
- Ryan, R. M., & Deci, E. L. (2007). Active human nature: Self-determination theory and the promotion and maintenance of sport, exercise and health. In M. Hagger & N. Chatzis (Eds.), *Intrinsic motivation and self-determination in exercise and sport* (pp. 1–19). Champaign: Human Kinetics.
- Sagar, S. S., Busch, B. K., & Jowett, S. (2010). Success and failure, fear of failure, and coping responses of adolescent academy football players. *Journal of Applied Sport Psychology*, 22, 213–230. <https://doi.org/10.1080/10413201003664962>.
- Sarrazin, P., Vallerand, R., Guillet, E., Pelletier, L. G., & Cury, F. (2002). Motivation and dropout in female handballers: a 21-month prospective study. *European Journal of Social Psychology*, 32, 395–418. <https://doi.org/10.1002/ejsp.98>.
- Scanlan, T. K., Carpenter, P. J., Simons, J. P., Schmidt, G. W., & Keeler, B. (1993). An introduction to the Sport Commitment Model. *Journal of Sport and Exercise Psychology*, 15, 1–15. <https://doi.org/10.1123/jsep.15.1.1>.
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74.
- Schultheiss, O. C., & Pang, J. S. (2010). Measuring implicit motives. In R. W. Robins, R. C. Fraley & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 322–344). New York, London: Guilford.
- Smith, G. T., Fischer, S., & Fister, S. M. (2003). Incremental validity principles in test construction. *Psychological Assessment*, 15(4), 467–477.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55.
- Unierzyski, P. (2003). Level of achievement motivation of young tennis players and their future progress. *Journal of Sport Science and Medicine*, 2, 184–186.
- Vaeyens, R., Lenoir, M., Williams, M., & Philippaerts, R. M. (2008). Talent identification and development programmes in sport: current models and future directions. *Sports Medicine*, 38(9), 703–714.
- Wenhold, F., Elbe, A.-M., & Beckmann, J. (2008). *VKS: Fragebogen zur Erfassung der Volitionaler Komponenten im Sport* [Questionnaire for recording the volitional components in sports]. http://www.bisp-sportpsychologie.de/SpoPsy/DE/Diagnostikportal/Motivation/Sportlerfragebogen/vqs_2008/vks_einfuehrung.html?nn=3014646. Last access: June 24, 2019
- Zibung, M., Zuber, C., & Conzelmann, A. (2016). The motor subsystem as a predictor of success in young football talents: a person-oriented

study. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0161049>.

Zuber, C., & Conzelmann, A. (2014). The impact of the achievement motive on athletic performance in adolescent football players. *European Journal of Sport Science*, *14*, 475–483. <https://doi.org/10.1080/17461391.2013.837513>.

Zuber, C., Zibung, M., & Conzelmann, A. (2015). Motivational patterns as an instrument for predicting success in promising young football players. *Journal of Sports Sciences*, *33*(2), 160–168. <https://doi.org/10.1080/02640414.2014.92882>.