SPORT PSYCHOMETRY, PARTICIPANTS AND INVARIANCE: A CRITICAL REVIEW

PSICOMETRIA ESPORTIVA, CARACTERIZAÇÃO DOS PARTICIPANTES E INVARIÂNCIA: UMA REVISÃO CRÍTICA

Flávio Rebustini1,4, Marcos Alencar Abade Balbinotti2, Renata Eloah de Lucena Ferretti-Rebustini3 e Afonso Antonio Machado4

1Universidade Estácio de Sá, São Paulo-SP, Brasil.
2Université du Québec à Trois Rivières, Canadá.
3Universidade de São Paulo, São Paulo-SP, Brasil.
4Universidade Estadual Paulista, Rio Claro-SP, Brasil.

RESUMO
Este artigo apresenta uma revisão crítica de como as variáveis do contexto esportivo e dos participantes podem afetar a busca por evidências de validade dos instrumentos de medida. Apesar da literatura apontar que fatores como sexo, idade, tempo de prática, tempo de experiência, experiências esportivas, fases e regimes de treinamento e competição, nível socioeconômico e educacional, dentre outros têm um efeito importante nas variáveis psicológicas, pouco tem se explorado sobre tal efeito. Recorrentemente, os estudos têm reportado basicamente a idade, sexo e modalidade esportiva. A ausência do controle dessas informações sobre o contexto e o participante pode provocar imprecisão das medidas, limitação da aplicação, erros de interpretação e intervenções inadequadas, perda da reprodutibilidade e limitação de estudos comparativos. Além disso, impossibilita a testagem da invariância. Desta forma, torna-se premente e mandatório um rigor maior sobre a gama de indicadores que caracterizam os participantes e o contexto dos estudos no esporte.


ABSTRACT
This article presents a critical review of how sporting context variables and participants may affect the validity evidence of measurement instruments. Despite the literature presenting, there is significant effect of gender, age, practice time, length of experience, sporting experience, stages, training loads, competition, as well as socioeconomic and educational level, among other psychological variables, little has been explored about such effects. Recurrent studies have basically reported age, gender and sport. The absence of control of this information in the sporting context and the participant can cause measurement inaccuracy, limitation in application, errors in interpretations and inappropriate interventions, loss of reproducibility and limitations in comparative studies. Furthermore, it prevents analysis of invariance, compromising their stability. Thus, it is urgent and mandatory to tighten the range of indicators that characterize the participants and the context of studies in sports.


Introduction

Psychosocial instruments in sport have been frequently used by professionals from several areas. In view of this fact, the first question which has to be answered is: do the instruments which have been used address adequate quality requirements and techniques in the search for evidence of validity? First, it is necessary to point out the concept of validity. Thus, Coaley1 points to the definition of validity made by the British Psychological Society as one of the most rigorous, which understands that "validity is the extent to which a test measures what it intended to measure, and the extent to which it is possible to make appropriate inference from a test". The second aspect contained in the question can be answered by the findings of Silva et al.2. The researchers verified sport instruments, and in
their conclusions they emphasized the scarcity of instruments validated in Brazil for the sporting scenario, their low quality and the need to pay attention to the unique characteristics of each modality, age group, and level of performance, among others.

In this sense, Furr & Bacharach claim that the tests should be able to compare the behavior of different people (interindividual differences) or the behavior of the same person at different points in time and under different circumstances (intra-individual differences). Thus, the psychometric quality of a test has implications on the interpretations, meanings and resulting interventions. This is reinforced by the recommendations of the International Test Commission regarding the need for efficiency, precision, accuracy of scores, analysis and reporting of the testing process. In order for these postulates to be achieved, a detailed characterization of the participants at all stages of an instrument development and of the techniques used in data collection (instruments and the form of development) is essential. This article focuses on this minute specificity of who the participants are and their characteristics, as well as the environment and the moment when data is collected. We must emphasize that these variables regarding the individuals in context definitively affect the stability and efficiency of the instrument and will be analyzed throughout this article.

Before going further into these instrument stability and invariance questions, we must go over the characterization of the subject and the environment. In view of these considerations, the objective of this article is to perform a critical reading of how specific factors of the sport can affect the search for evidence of validation of sporting instruments and how they may affect data from psychological testing.

Characterization of the participants and the sporting context

Despite books serving as a framework for research methods on the sports environment pointing to and recommending some of the characteristics of the participants and the context, it does not appear to have been adequately explored in studies seeking evidence of validity of psychosocial instruments. To reinforce this need, Takahashi points out that to examine the demands of the sport, it is necessary to discuss the interaction between the body and the context or its surroundings. For Thomas, Nelson & Silverman, some characteristics are essential and should be collected and analyzed in the context of sport and physical activity: age, gender, training level, performance level, size, and specific types (athletes, cyclists, runners). Turato suggests the need to stratify the participants’ composition in scientific research, and adds other variables: schooling/education level, by ethnicity and/or origin, by family role, profession or occupation, health condition, and religiousness, among others. Taking into account only these few variables as reference, it would already be possible to identify the weakness of studies in describing who and what we are researching, greatly compromising the study reproducibility, weakening the comparison between the participants and context variables. Consequently, we can indicate interventions - in the broader sense of the term – namely mitigations and solutions based on insufficient and weak results and interpretations. It should be noted that any minor change or inadequate intervention can be disastrous for the athlete’s or team’s performance, since competition sports require very high levels of precision and results.

For Sparkes & Smith, the researcher must understand the particularities of the context that will be researched, and the influence of this context on the participants' thoughts, beliefs and actions. Thus, the researcher should describe the participants and their life circumstances with details that allow the reader to evaluate the relevance and applicability of the study results. According to Salazar, Crosby & DiClemente, the purpose of any sampling technique is to maximize sample generalization to the population. In this way, we cannot extrapolate research findings to the population if there is not an adequate characterization of
the group under study.

We can look at the characterization issue from another angle. There are several articles and scientific studies that address psychological changes resulting from sports\textsuperscript{12,13}. However, we do not give emphasis to the fact that these changes may conversely interfere in interpreting the results and data of an instrument, when detailed information of the individuals and the context are not collected. As demonstrated by Schaal et al.\textsuperscript{14}, epidemiological studies focused on mental health in the sporting environment are very rare. In their extensive study they conducted with 2067 high-performance French athletes, the authors found: a) significant differences between men and women, where women were more likely to be diagnosed with a psychological problem and more likely to cope with the difficulties encountered in the sporting environment; b) that there are important differences between sports modalities; c) that the demands and pressures of practicing a specific sport can act in a significant way as a socioenvironmental risk factor, and that when combined with personality and genetic predispositions, they could facilitate the development of some disorders.

Phillips et al.\textsuperscript{15} stress that a theoretical approach for performance studies in sport, expertise and talent development must be focused on a dynamic and complex system. With a different focus, Dunwoody\textsuperscript{16} pointed out that psychologists should look beyond the individual, suggesting that they have neglected the role of the environment and have focused on the process and structure of the body in an isolated manner. Even more bluntly, Anshel & Lidor\textsuperscript{17} questioned studies that use psychological measures for detecting sports talent. It is important to note that the questions raised by the authors are eminently based on conceptual and methodological questions, such as: psychological measures that have low validity, based on flawed research procedures; lacking consistency in defining elite versus non-elite athletes, with inherent bias in the athlete selection process, and containing inaccuracies in measuring sports personality, among other concerns. Next, we analyze some of the characteristics of participants and contexts.

\textit{Age and gender of the participants}

In an extensive study\textsuperscript{18} that monitored 82 events of 5 modalities (athletics, swimming, weightlifting, road cycling and speed skating), the authors found a stability in the difference in outcomes between men and women of about 10% since 1983. They conclude that after a significant reduction in the disparity of sports results between both genders before 1983, men and women’s results have developed in parallel and in the same direction. In addition to performance issues, it is necessary to highlight that there are important psychosocial and morphofunctional differences between male and female athletes that should be considered in describing the research subjects.

Another aspect that is part of the sports environment that has not been described in evaluating studies and developing measurement instruments in sports and physical activity is the relative age effect (RAE) on sports development stages. There is a tendency for coaches who work with young athletes to select those who were predominantly born in the first quarter of the year (January through March) because they understand and assume that these youngsters are more physically fit and therefore tend to be stronger, faster and more skillful. For Baker et al.\textsuperscript{19}, relative age is an arbitrary selection process associated with disadvantages as to the participant's age of birth.

The first study on this issue\textsuperscript{20} identified that the number of athletes born in the first quarter of the year was 4 times greater than those born in the last quarter in a Hockey League. They also found that the dropout rate of athletes born in the last quarter of the year was much higher than those born in the first quarter. In more recent studies\textsuperscript{21,22}, relative age continues to play an important role in sporting neglect of young athletes. RAE has created a vicious cycle,
coaches look for athletes born in the first months of the year aiming for a physical advantage over the others, which prevents a normal distribution of birth months when compared to the population of non-athletes. There is no doubt that if this issue is not controlled, especially in studies comparing young athletes with young populations, mistaken analyzes and conclusions will be produced. This is because there will be a trend of better results among young athletes, because they are older than the non-athlete population distribution. RAE practically makes anthropometric and motor development studies unfeasible if this variable is not controlled, in addition to raising doubts about how the psychological variables behave when we control for relative age.

Hancock, Adler & Coté propose an explanatory model for the effects of relative age and emphasize that they do not exclude the impact of issues related to maturing, physical development and cognitive abilities, but they understand that the heart of the problem lies in the way parents, coaches and athletes interpret these mechanisms, meaning that this is what really creates the RAE. This encompasses psychosocial aspects and aspects related to the functioning of sport environments as being sensitive to RAE.

If relative age is one of the factors responsible for sports abandonment, by the same logic psychosocial aspects are present. In the systematic review developed by Crane & Temple on sporting abandonment, 5 major themes were detected: lack of fun, perception of competence, social pressures, competing priorities, and physical factors. Of the 5, only 1 (physical factors) is not associated with psychosocial issues. So, why do studies on psychosocial sporting assessments not bring relative age information into the socio-demographic description of the studied groups? In other words, solely using the year of birth as a reference is not enough; and that the period of the year when the athlete was born is also essential.

Sport modality, practice time, experience time and sporting experience

At times we encounter studies that do not distinguish the sporting modalities of the participants, mainly in studies that deal with the instrument development. This occurs more frequently in studies with psychological evaluations. Let’s consider the evaluation of fear in sport; is it possible compare athletes of extreme sports with ordinary regular athletes without considering the context in which they are inserted? Brymer & Schweitzer emphasize that in spite of the intense sensation of fear that extreme athletes feel, they still do not allow themselves to be immobilized nor stop practicing the sport. How can we accurately assess fear if we fail to indicate how the environment and practice act on the athlete?

There is little doubt that extensive sports experience is needed to increase performance as well as to achieve higher levels of performance. According to Birrer & Morgan, the difference between professional and semi-professional athletes and between winners and losers has become progressively narrower. This approach between professional and semi-professional athletes, however, does not exclude the significant difference between levels of performance.

Characterizing the sporting modality cannot exclusively explain the psychological assessment results, but it cannot be ruled out. In this way, competitive experience comes into play. Mellalieu, Hanton & O’Brien and Fernandes et al emphasize that the type of modality and the level of competitive experience influence the interpretation of the symptoms usually experienced in pressure situations. We can transpose this idea into people who practice physical activity who are not associated with performance sport. Childs & Wit found that physically active adults are more resistant to the emotional effects of acute stress. It is also possible to find differences in mood in high intensity exercises when comparing to the previous experience of practicing sports.
However, there are some peculiarities when focusing on sports experience, where assessing the competitive level only by years of experience is not enough. In competitive sport with young athletes, it is very common for athletes to also take part/perform in other categories; that is, an athlete who demonstrates better performance conditions in the youth category will usually also participate in the older child-juvenile category, not only during matches, but also regularly training on both teams. In some modalities, child and youth athletes perform in their own category, in the juvenile as well, and sometimes even train and compete with adults. That said, it is clear that athletes with the same amount of experience/time practicing sports have significantly different experience, intensities and training complexities.

Recently, Swann, Moran & Piggott proposed a classification model of the athlete's athletic level for research in sports psychology composed of 5 variables: Performance level; Success level; Experience level; Competitiveness level in the country; and Global competitiveness level (international competitions). The authors designed a way to classify athletes into 4 levels: semi-elite, competitive-elite, successful-elite, and world-class elite athletes. In this scenario, if experience affects psychological variables, we can understand that the higher the level of competition, the greater the chances of finding athletes with better physical, technical and tactical abilities, as well as psychological.

We can highlight the need to know about the level of competition of the athlete due to the effects this has on the psychological variables. A meta-analysis study on cognition and sporting expertise found that high-performance athletes/experts have better results on declarative memory tests, attention, attention allocation, improved perception and information, anticipation, and decision-making skills.

It is crucial to understand that top athletes usually have interdisciplinary technical committees, especially in the world's elite sports countries where these commissions have sports psychologists, and therefore athletes tend to be better psychologically prepared than those who do not have this support. Harmison stresses that for athletes to achieve their development and optimal performance, psychological preparation plays a vital role. If improving psychological abilities increases performance, then why when we characterize the participants do we not ask if they have meetings/coaching by a specialist in sports psychology and/or do they train their psychological abilities? And even further, is this preparation systematized?

Balague emphasizes the importance of psychological periodization and the need to consider a series of implications in the training model that corroborates with what has been exposed so far:

a) Different levels of sporting ability and experience may have a greater impact on the sequence of skills to be taught;

b) Different sports have different requirements/demands. Particularly, collective sports require different elements than individual sports, both in timing and scheduling, and due to environmental differences;

c) This model implies that psychological skill training should be performed in conjunction with physical training, in the gym, at the track and/or field, instead of at the psychologist's office. Co-operation between psychologist and coach is essential, and many of the interventions will be performed by the coach with the psychologist acting as a consultant.

d) The psychologist working within this framework needs solid knowledge regarding sports and sport science.

These points are reinforced by the study of Lidor, Blumenstein & Tenenbaum carried out with European basketball athletes, and by Holliday et al. in periodizing mental training. In this way, there is an advantage for the athletes who count on this support;
consequently, it is very likely that they will achieve better levels on psychological tests than those who do not rely on systematic psychological work and experience sports in a different way. An analysis of the variables on this topic exposes the need for greater care in reporting them in the research, given the evidence that they can significantly interfere in the study results.

*Sports training phases and regimens, and competition*

Hanin indicates that there are three criteria that can be used to evaluate performance results: group, interindividual and intra-individual. In addition, effects of periods and training regimens on performance and psychological variables are notorious.

It is not uncommon to find sports studies that do not mention whether data collection was conducted during the training or game period, especially characterizing the type of competition (dispute level and form). At first it may seem to be superficial information, however, when looking closely at some competition models, we must immediately deny this superficiality and even the simplicity of the fact. So many collective and individual sports have various forms of competition. In individual sports, it is common for competitions to last for a whole day or even two or three days. In the case of swimming, it is not uncommon that the athlete competes in several events on the same day. This occurs quite often in quadrangular finals in team sports. Tournaments in which athletes play in a quadrangular system on the weekends are common in American university sports. Even in the biggest tournaments in the world (world championships and Olympics), it is not uncommon to compete several days in a row.

Added to the fact that many modalities have very peculiar training regimens and games in which the athletes train in the period before the game, or if they compete in the morning, they go to training soon after the competition. In this way, it can be expected that the results of collections during training and match periods lead to differences. There are several studies dealing with emotional changes in training and games. Anxiety studies, for example, have constantly analyzed variation in athletes’ levels in a pre-competitive and competitive phase. Interdonato, Miarka & Franchini indicate that anxiety in competition is significantly higher than in training for both men and women, and it is significantly higher for women than men.

A point that lacks the integration of professionals working with sport psychology and sports evaluation is the field of sports planning and the effects of loads and training regimens on emotional aspects. There is an extensive list of symptoms related to overtraining and lack of adequate recovery from training loads and many of them are emotional aspects such as: depressing mood, generalized apathy, low self-esteem, emotional instability, agitation, and irritability, among others.

Literature has brought forth several studies that deal with the effects of training intensities and competitions on psychological and emotional states. Parry et al. found significant differences in anxiety, mood, and perceived exertion before and after events among Ironman athletes. Rebustini et al. found significant differences in mood with the variation of training loads in volleyball players, and there are similar results for soccer athletes and on the self-confidence in rhythmic gymnastics athletes. The high intensity of training brings with it an important variable, namely pain. Pain is part of the routine of athletes and it does not have to be due to an injury; it stems from the effort to reach very high levels of performance, which as a consequence will cause emotional changes in the athletes.

Several articles have addressed the effects of physical activity on anxiety, depression, mood and stress sensitivity. Recently, Jung, Bourne & Little verified affective response
before, during and after exercise sessions of three exercise regimens. Significant changes in self-efficacy and fun between different types of activities were detected. In a similar study, Saanijoli et al. indicated that high-intensity interval exercises presented a significant increase during and after physical activity in sedentary individuals.

Therefore, regardless of whether it is in high performance athletes (training and competitions), regular physical activity practitioners or beginners in sports, loads and exercise intensity significantly affect emotional states. A lack of this information prevents a precise understanding of the test results, which has an effect on the context of data collection, since this depends on detail of the environment that the individual is inserted in order to be accurately understood. For example, if a hypothetical set of athlete data was collected during the final preparation phase (championship finals or major competitions), what would generally lead athletes to their best conditions? If another hypothetical set of data was collected earlier in the season when athletes are out of shape/practice, is the analogy the same between a regular practitioner and a sedentary individual? Can we effectively say that one group of athletes has better test scores than the other group if this condition is not presented? This indicates that if the invariance is not tested for such interferences, we will most likely have unstable and imprecise instruments.

The lack of information on the regimens, the intensity of the physical training and activities imposed on the athletes, and on the practice of physical activity, as well as the moment in which the data collection was applied seems to us to be a limitation in the analysis.

**Socioeconomic and educational level**

The non-insertion of the socioeconomic level in the characterization of the participants can be considered incomprehensible. In a study on the socioeconomic level of children, Vandesdriessche et al. found that a low socioeconomic level is associated with reduced sports participation and membership in clubs. In addition, low and medium socioeconomic levels are associated with higher body mass, body mass index and fat percentage, and also with poor motor coordination, especially among girls. Higher socioeconomic levels tend to be more physically active; however, 42% of the studies were not significant or showed a negative correlation. In any case, an important variable is pointed out since these oscillations may be linked to the cultural issues of each region or country. The same author sought to identify whether socioeconomic level affects the anxiety of the athletes; the results indicate that athletes with higher socioeconomic level presented higher cognitive anxiety and self-confidence, and lower somatic anxiety than athletes with average socioeconomic status. Although the results were not significantly different, in this case they were limited by the number of study participants. At any rate, there seems to be a trend that needs to be better understood.

It is possible to easily notice the absence of information regarding the educational level in studies involving athlete populations. However, association policies of physical activity and education are notorious and extensively documented outside Brazil. Sports organizations in Brazil are not structured to give educational support to high-performance athletes, despite the importance and effect caused by a double career. In an extensive study of French, Finnish and English athletes, it was possible to detect that their educational experience facilitated sports development, and skills developed in the sporting context were seen as transferable for education and the work environment, while also listing a number of benefits of the association between study and sport, such as: a) planning and organization of competencies, teamwork; b) interpersonal skills; c) ability to set goals, prioritize and monitor achievements; d) maintaining commitment; e) proficiency development and analytical skills which can facilitate communication in contexts with a technician; f) leadership and; g)
awareness of networking and business. In this context, the effect of education on sports and vice versa is evident. Another effect is the development of a double career of transnational athletes, who receive scholarships to study and compete away from their home countries, where there is also a lack of professionals who can support athletes with a dual career. Also, the importance of the athlete's cognitive development in tactical learning cannot be understated, which is nothing more than the logic of game, action and counteraction.

If this is indeed the case, why do sports assessment studies routinely fail to mention athletes' educational levels? What's more, does this educational level not interfere with the way athletes respond to assessments? This aspect is so noteworthy that the Item Response Theory (IRT) considers the difficulty of responding to an item in its theoretical framework.

**Sports Psychometrics and Psychological Testing in Sport**

To conclude, we propose a reflection on the variables that have been discussed in this text: suppose we applied a test or a set of instruments to a group of athletes, and we took it to two groups of psychologists or specialists in sports psychology who had no contact with the athletes; for one of them, we would provide all the variables and characteristic information of the athletes and the sports environment in which they are inserted; while for the other professional, we would only present the psychological testing results. There is only one question remaining: Would the interpretation and its resulting actions be the same among the professional groups? Most likely they would not.

For Urbina, a psychological test is a systematic procedure for obtaining behavior samples relevant to cognition, affectivity and interpersonal functioning, and to score and evaluate these samples according to the standards.” Moreover, Furr & Bacharach claim that there are three important components to consider in developing a test: 1) the test involves some behavioral sample; 2) the behavioral sample must be collected in some systematic way and; 3) the test’s purpose is to compare the behaviors of two or more people. Associating the information of these authors, the comparison between athletes for whom their precise information and the context in which they participate were not collected, and therefore the comparison becomes fragile. Going back to the basics of psychological testing, “the place and time at which the test was applied, as well as the circumstances under which it was filled or the mode of administration used, may affect its results.”

Another sensitive point is the impossibility of guaranteeing the instrument's invariance, meaning the degree to which an instrument’s configuration and parameters are equivalent for different groups of people. It is precisely the concept of invariance that sustains the observations made throughout this article. It should be noted that invariance is an analysis that has gained relevance in recent years. For Brown, “Factorial invariance is an important aspect of test development. If a test is intended to be administered in a heterogeneous population, it should be established that its measurement properties are equivalent in subgroups of the population (e.g. gender, race), and they add that a test is said to be biased when some of its items do not measure the underlying construct in a comparable way across all groups. It can be affirmed that we cannot consider the population of athletes and practitioners of physical activity or modalities as homogeneous. Consequently, performing analysis of invariance becomes mandatory, and therefore characteristics of the individuals who participated in the collection in search of validating the instrument together with the collection context are fundamental points. Furthermore, they alone sustain the questions made throughout this article, and this is added to the fact that the literature rarely focuses on temporal invariance, which is intended to longitudinally measure the invariance of the test, thus measuring the temporal changes of the instrument.

According to Damasio, the assumption of invariance responds to some points,
namely that: a) the factorial structure of the instrument is the same in different groups, b) the items that make up one factor and the instrument have the same importance for different groups; c) the scores of one group can be compared to other groups; d) the items present errors of similar measures for different groups; e) the level of variance between the factors differ between groups, and; f) the covariance between the factors is the same between the groups. Given this set of aspects, is it possible to guarantee the invariance of an instrument without the information and variables analyzed in this article?

Several subgroups have been analyzed in instrument development studies: athletes, physically active people, and fans; gender and sport modality; gender and level of physical activity; and language, gender, age and type of sport. However, it was not possible to find articles related to sports and physical activity that deal with a more important range of variables and which allowed for a detailed description of the study participants’ collection environment, which is evidently an inconsistency when analyzing the rigor and precision of a study. Obviously, analysis of a larger number of variables leads to the need to work with a larger number of participants, allowing for an application of more robust instrument validation techniques. In any case, invariance analysis in psychometric studies is mandatory, and in order to be performed, characteristics of the participants and the context to be used in testing must also be collected.

Conclusions

Reflections and criticisms throughout this article lead us to a series of other reflections on how accurate and robust the instruments applied to sport are, in addition to the consequences of its use for intervention and daily work when based on research results with uncontrolled limitations. In general, it is possible to affirm that a precise description of the participants and the sporting context have been a focus of immense fragility in sports psychometry, and can lead to some problems: imprecision of the measurements, limitation of the instruments’ application, interpretations and consequently inadequate interventions, loss of reproducibility of the procedures and a limitation of comparative studies. This analysis provokes the urgent and mandatory need to review how researchers are reporting the variables that characterize participants and their contexts in sport.

References


32. Rebustini, F. Interferência dos fatores externos sobre os estados de humor em jovens atletas de voleibol. [Dissertação de mestrado em Ciências da Motricidade]. Universidade Estadual Paulista, Instituto de Biociências de Rio Claro; 2005.


Address for correspondence: Flávio Rebustini. LEPESPE. Av. 24ª, 1515. Rio Claro, SP, Brazil. CEP: 13506-900. E-mail: frebustini@uol.com.br