Routledge Handbook of Applied Sport Psychology
A comprehensive guide for students and practitioners
Edited by Stephanie J. Hanrahan and Mark B. Andersen
Test use in sport psychology practice can certainly be controversial. The widely discussed credulous versus skeptical debate (Morgan, 1980) regarding the efficacy of psychological test use in sport psychology had a significant effect on the field. Sport psychologists who were professionally active in the 1970s will recall the earnest and somewhat divisive discussions. Perhaps as a result of these discussions sport psychologists gravitated toward becoming either staunch advocates (i.e., credulous) or staunch adversaries (i.e., skeptical) of testing. For example, the credulous side believe that personality data derived from questionnaires are extremely useful in predicting success, whereas the skeptical side believe questionnaire data have little relevance and usefulness (cf. Morgan, 1980). Now, some 35 years after the fervor of the credulous–skeptical debate, we can reflect that both sides were using data derived from somewhat dubious methods based on current standards and arguments. Also, in reviewing the relevant literature from the period, there were many inflammatory comments and statements proposed from both the skeptical and the credulous.

Later, when The Sport Psychologist published two special issues on consultants working with professional and Olympic athletes (Roberts, 1989; Roberts & Halliwell, 1990), forthright opinions were again expressed that echoed the earlier debates. In more recent times, the development and psychometric evaluation of standardized tests in sport psychology has remained highly visible in the published literature, but the debate over the use of testing has waned. Nevertheless, readers who have entered the field since the early 1990s might be unaware of the credulous–skeptical debate. Sport psychologists who were active participants in the debate may still harbor attitudes about testing, informed largely from a sport psychology landscape that bears only some resemblance to that of the present time. I have no intention, however, of reviving debate about the relative merits of testing here, but rather to engender an applied practitioner perspective of test use and relevant considerations. Nevertheless, the credulous–skeptical debate provides an historical backdrop to current use and non-use of psychometric instruments in applied sport psychology service delivery.

Perhaps as a result of the fervor of earlier debate, the subsequent decades of psychological testing discussions have, in my view, retained an austere quality largely conducted within the walls of academia. Sport psychologists nowadays are sometimes asked to provide testing
for talent identification and athlete profiling purposes. Some sport psychologists can be adamant about not using tests, or alternatively, professionally irresponsible in using tests without due diligence. There is certainly a delicate balance to be sought when using tests in applied contexts. I encourage practitioners who choose to adopt psychological testing as part of their services to reflect on the issues presented here. Graduate students and early-career applied sport psychologists should at least find some “grist for the mill” in answering practical questions that have largely been ignored in the published literature in terms of informing professional practice.

**Testing times: reflections on test development and use in sport psychology**

After years of research in sport psychology, the number of sport-specific tests available is now rather daunting (see Ostrow, 2002). Although the quantity of tests has grown substantially over the years, the quality of tests is varied. Even though the following list is deliberately cynical, experienced test users might recognize these types of instruments:

- **The pre-statistical revolution test** – Developed in an era when statistical methods were less sophisticated and the spotlight of examination less bright. These tests that were generally developed in good faith and based on accepted test construction practices have sometimes been found, after reexamination, to be fundamentally flawed.
- **The unsuccessful prototype test** – Developed specifically for doctoral studies and usually accompanied by the following lamentation, “If only I had known how difficult it would be to develop a test”.
- **The model of propriety test** – Developed using modern standards in statistical analysis (e.g., confirmatory factor analysis). Yet the robust statistical indicators of such tests belie the questionable repetition of items that ensures excellent internal consistency at the expense of a decent representation of the construct, and in turn, diminishes the practical usefulness or external validity of the test.
- **The search for the Holy Grail test** – Developed to identify or unlock the psyche of sporting greatness. Fortunately, the lure to develop such tests generally hooks those from outside or on the fringes of sport psychology, and clearly reflects a degree of naïveté.
- **The open to deception test** – Developed for a specific purpose but with little thought given to social desirability issues and self-deception (where the respondents actually believe their positive self-reports) and impression management (where the respondent consciously obscures the truth). Currently, many test respondents are relatively savvy and test-item transparency is becoming a highly relevant issue depending on the type and purpose of the test.
- **The top-down theoretically derived test** – Developed with careful consideration of research findings but with little relevance, sensitivity, and hence, practical usefulness in consulting situations.

Admittedly, there are many well-constructed tests in sport psychology. Excuse the analogy, but at times it seems that test construction is akin to an uncleared minefield where treading carefully might simply not be enough. Many sport psychologists have experienced the scorched earth of developing tests that probably seemed like good ideas at the time: I certainly have! The point is that psychological testing is a serious business, and to extend the minefield metaphor,
the soldiers might appreciate a little forewarning about the dangers that lie ahead. Clearly, sport psychology as a field needs to provide our “soldiers” (i.e., applied sport psychologists) with the most advanced and sophisticated weaponry. OK, enough of that metaphor!

Collectively, do we feel that the psychological tests available for use by sport psychologists represent best practice tools? The question needs to be asked because the expectations from elite sport organizations that contract sport psychologists are often quite lofty in terms of what they believe practitioners can deliver. Common questions sound something like, “what tests do you have that will enable us to identify the most promising athletes in advance or those with good character and attitude?” In this context of high expectations and athletes becoming better educated about mental skills and psychology generally, the instruments we use need to be of the highest quality. Historically, sport-psychology-specific tests have been developed in university settings, primarily for research purposes. Although a number of tests have been constructed that are adaptable to both research and applied settings, there is certainly scope for further development of tests that reflect the specific needs of applied sport psychology practitioners. The challenge is to develop tests that are market driven and remain reliable while being ecologically valid. Andersen, McCullagh, and Wilson (2007) have mounted a case for a stronger focus on calibrating tests against real-world behaviors. Arbitrary measurement of psychological constructs is relatively common in sport psychology and can be misleading (e.g., does a score of 4 on an anxiety test item imply an athlete is twice as anxious as an athlete scoring 2 on the same item?). Furthermore, and along the lines of Andersen et al.’s concerns, what do scores on psychological inventories tell us about real-world performance or behavior? Andersen et al. argued that when sport psychology tests are calibrated against real world behaviors, the metric or evidence is more trustworthy and hence legitimizes the testing enterprise. In reality, few psychometrics in sport psychology research and practice are calibrated against the real-world variables of interest (e.g., performance).

Understanding fundamentals: a psychometric mindset

Included in the training for most psychologists is a course or two on assessment/psychological testing. These courses are invaluable in providing foundations about the broader parent discipline of psychological assessment, and specifically, test development, administration, and reporting. Before moving to developing a “psychometric mindset,” it would be remiss not to recall that quantitative psychological tests represent just one class of technique in psychological assessment. Well-rounded assessments involve multiple sources of information such as interviewing and observation in addition to psychological testing. Some psychologists (e.g., Lines, Schwartzman, Tkachuk, Leslie-Toogood, & Martin, 1999) advocate the use of sport-specific behavioral assessments such as a type of checklist that reflects the specific technical language of the sport (e.g., on backstroke: “Head still,” “Hips high”).

Although there are thousands of psychological tests and hundreds of sport psychology tests available, many of them are unsuitable for applied settings. The limitations of tests are many and varied, but can generally be traced back to limitations in reliability or validity. Familiarity with the foundations of psychometrics assists sport psychologists in making informed choices about which tests to use, for what populations, and in what contexts. Because tests do not get published if they are not reliable, test developers often pay close attention to assessing and demonstrating reliability. Beyond basic inter-item and test-retest reliability, various forms of validity testing such as factor analyses (or principle components
analysis) are minimum requirements. From a psychometric viewpoint, the current gold standard in terms of construct validity is confirmatory factor analyses. Although construct validity is well catered for with factor analytic techniques, test developers are sometimes not as diligent about the many other aspects of validity that require attention.

Few tests are initially published with systematic attention and coverage of: content validity, face validity, convergent/discriminant validity, criterion/predictive validity, and particularly internal/external validity. Yet from a practitioner’s perspective, validity questions are central. For example, does the test measure a characteristic that will carry over or transfer to the sports context? It may be years before all the relevant psychometric properties of a test have been examined. Tests that stay in vogue are characterized by some of the following features: resilience to peer-reviewed scrutiny; clearly defined uses and limitations; adequate administration, standardized scoring, and accepted reporting procedures; and, to a lesser extent, attractive presentation. Irrespective of whether a research or applied sport psychology career path is pursued, it is worth periodically scanning the latest psychometric studies regularly published in the sport psychology journals. Questions worth considering include: has the test been peer reviewed and published in a reputable journal? Beyond publication, how widely is the test used by other sport psychologists in research and/or applied work? Have other sport psychologists, apart from the test developers, published research pertaining to the psychometric qualities of the test? In what specific context could this test be useful in consulting?

Watch your step: suggestions for responsible and effective test use

Understanding the fundamentals of psychometric testing provides a reasonable starting point for the responsible use of tests in applied contexts. At this point, however, we are merely rounding first base. Before requesting that clients, who are generally coaches and athletes, take up their valuable time completing a psychological test, a number of issues require forward thinking. For efficiency, these issues are presented here as a series of questions akin to a checklist.

- First and foremost, for what particular purpose is the test designed?
- Is the test only available to registered psychologists, or is it in the public domain?
- Is there a cost associated with using the test, and what are the copyright limitations?
- Is there available normative data for target populations?
- What sport psychology colleagues (e.g., supervisor, professional mentor) have used the test and can provide you with professional guidance if necessary?
- What are the limitations of the test? Even the so-called gold-standard tests have limitations with which users need to be familiar.
- Have you completed the test yourself?
- Are you familiar with the nuances of the test and the administrative procedures?
- How should the test be introduced to the client?
- What type of reporting or feedback are you going to provide to the client?

Choosing tests: start with the end in mind

I expect that most readers are familiar with the rationale that for researchers the objective is to find the best fit between the research problem and the method chosen to investigate
the research question (i.e., elegant design). Similarly, in applied sport psychology contexts, tests/inventories/questionnaires are adopted with an expectation of better understanding a person, group, or presenting issue. By commencing with a clear view of what will be achieved by using a psychological test, we are at least envisaging the complete process from start to finish. Intentionality is the key. A broad overview of potential reasons for testing is a useful starting point. Possibly the most recognized reasons for using tests in applied settings is to measure susceptibilities, characteristics, tendencies, abilities, traits, states, and styles in relation to performance, training, mental skills, and topics frequently associated with applied sport psychology.

The term “diagnostic test” is broadly relevant to all tests but also carries clinical connotations. Making a clinical diagnosis of psychopathology may not necessarily form a substantial part of service delivery, or be part of how some sport psychologists conceptualize their work. Nevertheless, there are numerous studies that have demonstrated that significant psychopathology is present within athletic populations (see Brewer & Petrie, 2002). Irrespective of whether an applied sport psychologist has completed substantial clinical training, there is a professional responsibility to at least be familiar with prevalent clinical issues (e.g., personality disorders, eating disorders, substance abuse, anxiety disorders) and make referrals to suitably trained colleagues (e.g., clinically trained sport psychologist, clinical psychologist) if needed. We have moved well beyond the era of the educational sport psychologists who absent themselves from professional responsibilities regarding clinical recognition or awareness based on a limited clinical training. For identifying maladaptive behaviors and mental disorder symptomatology, a range of psychological tests are available that can be of assistance. For example, the Beck Depression Inventory (BDI-11; Beck, Steer, & Brown, 1996), a 21-item scale, guides respondents through an assessment of depressive symptoms over the past two weeks; the Eating Attitudes Test (EAT; Garner & Garfinkel, 1979), a 26-item scale, is a useful screening tool for clients who present with possible pathogenic eating; the Revised NEO Personality Inventory (NEO-PI-R; Costa & McCrae, 1992), despite primarily being designed to measure the big five personality traits, has a useful feature whereby the professional report generated with the NEO-PI-R software also includes a clinical diagnostic section that quantifies the likelihood of personality disorders (i.e., an Axis II diagnosis from the Diagnostic and Statistical Manual of Mental Disorders-IV-TR [American Psychiatric Association, 2000]). Where clients are shown to have an elevated pre-disposition to a particular personality disorder, the clinician could choose to carry out a follow-up assessment using a scale such as the Personality Disorder Interview-IV (PDI-IV; Widiger, Mangine, Corbitt, Ellis, & Thomas, 1995), which represents an in-depth clinical guide to assessing 10 personality disorders. Applied sport psychologists, depending on their level of clinical training and competence, will normally need to make judicious decisions about whether to refer or treat clients who present with clinical issues.

A great place to start: tests for developing insight

Postgraduate students often ask questions about psychometrics and express an interest in gaining proficiency in administering, scoring, and interpreting psychological tests. Apart from designated university psychometrics courses that students are generally required to complete as part of their training, how do students gain experience in psychometrics? Given the (often healthy) skepticism still prevailing in the field of sport psychology regarding test use, there is certainly value in working with tests to establish your own view of whether psychological tests are potentially useful (or not useful) in your professional
practice or research. Using selected tests and forming your own opinions is similar to reading the original primary work of Freud before forming opinions, as opposed to relying on secondary sources. A brief sketch of how to start will probably include, at the least, the following steps: (a) seeking supervision from an experienced psychologist; (b) reading the relevant current literature; (c) networking and talking to colleagues; (d) researching what tests are sometimes used with what populations and with what results; (e) establishing a clear purpose for administering a test while bearing in mind that tests are simply one assessment tool; (f) familiarizing yourself with local guidelines for ethical test use; (g) working with the client (e.g., consultation) or participant (research) to educate them on test use and establishing the reasons for testing, likely pros and cons, and responding to client questions; (g) administering, scoring, and interpreting the test; (h) possibly writing/producing a client report and debriefing the client; and finally, and importantly, (i) establishing a “where to from here?” That is, how will the test results inform ongoing work or benefit the client? See also Beckmann and Kellmann (2003) for a number of additional suggestions and perspectives on procedures for test use and feedback to clients.

A test use example: The NEO-PI-R

One of the few tests that I have regularly used with clients is the longer version of the NEO-PI-R (Costa & McCrae, 1992), which includes scales that measure susceptibility to anxiety, depression, and angry hostility within the neuroticism (emotional stability) factor. The NEO-PI-R is widely cited in the psychological literature. It has sound psychometric properties and is based on the modern 5-factor personality standard. The NEO-PI-R is not sport-specific and primarily serves three purposes within the sport psychology consultation context. First, it can be used to assist in establishing a relatively quick assessment into 30 personality characteristics and may facilitate initial intakes and help me come up to speed with the type of person with whom I am working. Second, its use may enhance insight by providing information that clients expand on in counseling sessions, and it may serve as a self-narrative primer. Third, although the NEO-PI-R is not sport-specific, it does provide a window into potential areas for future exploration and interventions and almost endless lines of behavioral inquiry that might later be observed, discussed, or connected to athletes’ sport and exercise involvement.

Another advantage of having an athlete complete a 5-factor personality measure such as the NEO-PI-R is the potential to better identify strengths and weaknesses in mental skills. For example, with some experience, combination scores on particular factors provide valuable clues that can then be followed up in further investigation in one-on-one work. A high anxiety–high self-consciousness combination might indicate a possible susceptibility to choking. A high angry hostility–low deliberation–high impulsiveness combination might indicate impulse control issues. High fantasy scores might indicate acuity for imagery. A high order–high achievement striving–high self-discipline combination might indicate perfectionist tendencies. The repeated use of a personality instrument, combined with working directly with athletes for an extended period, provides opportunities to test hunches about the potentially endless connections between personality profiles and observations of everyday behavior. In my consulting work with professional Australian Football League (AFL) players, I have found the NEO-PI-R to be exceptionally useful in the early stages of a consulting contract with AFL squads, which comprise close to 60 athletes and approximately five full-time coaches. Apart from enabling the consultant to identify potential clinical issues, a NEO-PI-R profile
will likely represent a useful personality overview of individuals with whom consultants will potentially be working over an extended period. Moreover, debriefs with clients, including access to their NEO-PI-R report, largely confirm in their verbal feedback that they find their personality profiles to be accurate. Occasionally, clients are almost astounded at how accurate they feel the reports are, and this feedback provides an opportunity to demystify the process by emphasizing that they are essentially “describing themselves through a questionnaire.” I must stress, however, that much of the above is contextual and based on personal experience, albeit over a 15-year period.

Debriefing with the client, if carried out sensitively, can improve the athlete’s self-awareness and insight. Goleman (1998) stated, “Self-awareness serves as an inner barometer, gauging whether what we are doing (or are about to do) is, indeed, worthwhile” (pp. 57–58). Giges (2004) encapsulated the interplay between self-awareness, sporting performance, and behavior:

Self-awareness involves knowledge of one’s own behavior, thoughts, feelings, needs, and wants ... A major advantage of such knowledge is an increased ability to be in charge of oneself, enabling individuals to be less distracted by their own feelings, wants, and values and better able to respond to others effectively. Self-awareness can also lead to improvement in sport performance. Without it, feelings can surface in unexpected or undesired ways. Consider coaches who ... to avoid feeling guilty about selecting poor strategy blame their athletes for not trying hard enough. Awareness of guilt might be difficult, however, for coaches who are unable to make an internal attribution and accept responsibility for their mistakes.

(p. 432)

**Cautionary tales: indentations, potholes, and craters**

Mistakes in the use of psychological testing can have relatively dire consequences. There isn’t the space here to provide an exhaustive list of potential pitfalls, but I will introduce a range of considerations that might be described as indentations (e.g., misinterpretation of an item) to craters (e.g., aggrieved clients angrily confronting a psychologist about their “negative profiles”). Potential pitfalls in using psychological tests generally fall into six broad categories: (a) choice of the test; (b) psychometric properties of tests; (c) administration and scoring errors; (d) interpretation errors; (e) reporting errors; and (f) safe-guarding profiles, scores, results, and associated reports. Choice errors can be avoided by being thorough in: assessing the psychometrics properties, determining the theoretical basis or construct validity of a test (not necessarily taking test developers’ promotional material or test retailers’ claims at face value), trialing tests with limited samples, investigating (where possible) tests within a domain that represents current best practice, subscribing to catalogs from reputable test providers, avoiding tests that are primarily geared for research purposes, and using other forms of assessment if you are uncomfortable with the quality of tests in a particular domain. Errors relating to psychometrics have been touched on earlier. To reiterate, however, although reliability is critical, applied sport psychologists are equally interested in various forms of validity. Administration and scoring errors are annoying because they can often be traced back to human error. Interpretation errors can result from a lack of in-depth knowledge about the test, under-interpretation/over-interpretation, limited sensitivity to sport context variables, poor debriefing skills, and misinterpretations from the client. In applied
settings, test use is unlikely to be a straight two-way exchange. Applied sport psychology environments are generally complex in delineating ownership of tests materials and best-practice procedures.

Understandably many sport psychologists are concerned about the ethical ramifications, the lack of predictive power of many tests, and the practical considerations in relation to athlete recruitment. Professional sporting organizations, will, however, argue their right to use psychological assessment tools to assist in minimizing risks, especially in terms of potential absorption of resources and the financial ramifications of poor selection decisions. Although familiarity with professionally drafted guidelines and procedures for test use is helpful, the specifics vary depending on the governing structure, local history of psychological accreditation, training requirements, psychological registration, and governing bodies. Furthermore, sport psychologists will inevitably be faced with challenging practical dilemmas and vexing questions that guidelines may not necessarily directly address.

Conclusions

In writing this chapter, I am acutely aware that we are touching on issues that warrant a more detailed and comprehensive discussion. Even with the sheer number of tests available, there is a notable absence of references to specific tests used in real-world consulting apart from a few examples showing how tests might be used for a particular purpose. The absence of real-world references is a reflection of the scarcity of dedicated contributions geared specifically to test use in applied settings. Perhaps this limited available literature on specific issues confronting applied sport psychologists who are contemplating using tests harks back to the history of tests in sport psychology being controversial. Irrespective of the issues raised here, psychological tests are merely tools that have the capacity to add both depth and breadth to assessment practices when used judiciously. Sport psychology remains a niche field and psychological testing a small sub-discipline. At present, the published sport psychology literature, in my view, is skewed toward psychological tests from an academic or researcher perspective. You may have heard the phrase “the psychologist is the instrument.” Thus, “the instruments might be more instrumental if better instruments were at their disposal.” See Box 12.1 for the main take-home messages from this chapter.

Box 12.1

Take-home messages about psychometric testing

- Attitudes to psychological testing in sport psychology have historically been informed by the credulous–skeptical debate.
- The quality of psychological tests developed in sport psychology varies considerably. Practitioners who intend to use tests are duty bound to choose psychometrically robust tests.
- The many challenges of using psychological tests and inventories responsibly should not be underestimated. Psychologists will often cite anecdotes or case examples of where test use has gone awry.
Sport psychologists will often be particularly interested in evidence of external validity and predictive validity when choosing tests in applied practice. Well designed psychological tests can provide extremely useful information about clients’ personalities, abilities, tendencies, and traits. Being clear about the purpose of using a test(s) in the early stages of a consultation is important for both the psychologist and the client. When using tests, careful attention to administration and scoring procedures, safe keeping of data, and report writing can circumvent potential problems. It is unlikely that psychologists will desist from continuing to employ tests and inventories. The challenge for the field of sport psychology is to develop high quality tests. That is, psychometrically sound tests that are tested for their own internal qualities.

References


