A call for policy guidance on psychometric testing in doping control in sport

Andrea Petróczi[1,2*], Susan H Backhouse[3‡], Vassilis Barkoukis[4‡], Ralf Brand[5‡], Anne-Marie Elbe[6‡], Lambros Lazuras[7‡], Fabio Lucidi[8‡]

[1] Kingston University London

[2] University of Sheffield

[3] Leeds Beckett University

[4] Aristotle University of Thessaloniki

[5] University of Potsdam

[6] University of Copenhagen

[7] University of Sheffield International Faculty

[8] Sapienza University Rome

* Corresponding author: Andrea Petróczi

Faculty of Science, Kingston University London, Penrhyn Road, Kingston upon Thames, Surrey, KT1 2EE, United Kingdom; phone/fax: +44(0)20 8417 2436; email: A.Petroczi@kingston.ac.uk

[‡]Listed in alphabetical order. All authors contributed equally.

Abstract

One of the fundamental challenges in anti-doping is identifying athletes who use, or are at risk of using, prohibited performance enhancing substances. The growing trend to employ a forensic approach to doping control aims to integrate information from social sciences (e.g., psychology of doping) into organised intelligence to accelerate the pursuit of clean sport. Beyond the foreseeable consequences of a positive identification as a doping user, this task is further complicated by the discrepancy between what constitutes a doping offence in the World Anti-Doping Code and operationalized in doping research. Whilst psychology plays an important role in developing our understanding of doping behaviour in order to inform intervention and prevention, its contribution to the array of doping diagnostic tools is still in its infancy. At the same time, we must acknowledge that socially desirable responding confounds self-reported psychometric test results. Further, the cognitive complexity surrounding test performance means that the response-time based measures and the lie detector tests for revealing concealed life-events (e.g., doping use) are prone to produce false or non-interpretable outcomes in field settings. Differences in social-cognitive characteristics of doping behaviour that are tested at group level (doping users vs. non-users) cannot be extrapolated to individuals; nor these psychometric measures used for individual diagnostics. In this paper, we present a position statement calling for policy guidance on appropriate use of psychometric assessments in the pursuit of clean sport. We argue that both self-reported and response-time based psychometric tests for doping have been designed, tested and validated to explore how athletes feel and think about doping in order to develop a better understanding of doping behaviour, not to establish evidence for doping. A false 'positive' psychological profile for doping (or even failing to produce a definite negative profile) affects not only the individual ‘clean’ athlete but also their

entourage, their organisation and sport itself. The proposed policy guidance aims to protect the global athletic community against social, ethical and legal consequences from potential misuse of psychological tests, including applications as forensic diagnostic tools in both practice and research.

Keywords: prohibited performance enhancement, athlete, drug, anti-doping, attitude, profiling, forensic diagnostics

Introduction

Owing to the recurring doping scandals, a degree of suspicion always falls upon competitive sport and its stakeholders. In a bid to assure the general public, athletes are looking for ways to pre-emptively prove their noble standing as 'clean athletes'. In recent years, athletes have made public pledges of support for global and national anti-doping campaigns such as the World Anti-Doping Agency's *Say NO to doping!* and UK Anti-Doping's *100%ME*. Individually, athletes are also taking ownership of the 'clean sport' heuristic, as exemplified by athlete Dee Dee Trotter who is using social networks to promote the assertion *'Test me, I am clean!'* Beyond anti-doping organisations, the independent not-for-profit organisation Bike Pure aims to promote clean cycling and has amassed a significant following.

However, high profile cases of prolonged and systematic doping, that have been retrospectively admitted or proven, cast a pall over any athlete's self-declared innocence. In a legal sense, one is innocent until proven guilty but in the public eye and the anti-doping sphere, this is not so much the case. Doping control builds on detection-based deterrence through doping testing, combined with education-based prevention. Whilst the latter encompasses all athletes under the auspices of the national/international anti-doping organisations and sport federations, the costs and logistics of drug testing prohibits the detection net to be cast far and wide. Consequently, routine measures to evidence clean status for a large number of athletes are not readily available. Periodically repeated analytical testing of all athletes' biological samples to continuously provide evidence for the clean status is not feasible for many reasons: (1) as argued above, it is not possible to evidence 'clean' status directly, only by the tacit assumption that all non-clean athletes are detected and removed; (2) the recently observed expansion of the prohibited substances, particularly with endogenous hormones and noble gases (e.g., xenon and

argon) poses an increasing challenge to detection-based doping control, (3) the cost is prohibitively high at an average of 300 US dollars for each routine test, with specialist tests being much more costly (personal communication, Olivier Rabin, January 18, 2013) and (4) management of such a system is not only resource intensive and inconvenient (Elbe, Melzer & Brand, 2012) and inherently paradoxical (Pitsch, 2013), but mandating such a system is also an infringement on athletes' human rights (Hanstad & Loland, 2009). The question is then how can one pre-emptively prove non-guilt?

Entrepreneurial initiatives appear to emerge as alternatives to analytical testing, reaching for cost effective psychometric methods readily available in the anti-doping researcher's tool-box. These tests are widely available (published or otherwise accessible), relatively inexpensive and non-invasive, with results easily stored and analysed if compared to any form of analytical tests based on bodily fluids and tissues. Although authoritative voices, such as WADA's former Chairman John Fahey advocates education - and thus social science approaches - over increased analytical testing effectiveness and capacity (Lane, 2014), financial investment has not followed such advocacy. The funding balance is still heavily weighted towards supporting the development of more sophisticated analytical techniques, rather than evidence-based prevention programmes (Backhouse, Patterson & McKenna, 2012). Despite this imbalance, recent years witnessed the emergence of new researchers and teams in the landscape of social science doping research. On the one hand, this expansion has had a positive effect on doping research by bringing diversity, variety and international flavours. On the other hand, it increases the risk of potential misuse of these psychometric tests and consequently, misinterpretation of the outcomes.

Our current concerns about the potential misuse of psychometric measures arose from a recent privately funded anti-doping initiative the Clean Protocol (http://cleanprotocol.org/), which aims to issue athletes with a ‘clean’ certificate upon a successful pass of a battery of psychometric tests, including a lie detector test. Despite that anti-doping organisations with sanctioning power distanced themselves from this initiative, some testing already took place on a voluntary basis. With indications for further testing mandated for teams, we felt the urgent need for informing end-users (athletes, entrepreneurs, anti-doping officials and researchers) about the limitations inherent in direct- and indirect psychometric measurements and issue a caution against employing these psychometric measures outside its intended use. Although the Clean Protocol is propagated as a positive approach by offering a ‘clean’ badge to those who can ‘prove’ that they do not use prohibited methods and substances (rather than identifying athletes who have doped), diagnostics do not work on this principle. It is the exact opposite. Because we cannot prove that something is absent, the initial assumption in any diagnostic procedure or statistical testing is that ‘something is not present’. This assumption then - if there is enough evidence to the contrary - is proven to be incorrect and thus rejected. Applying this position to sport, diagnostics tests (Clean Protocol included, along with any form of analytical doping testing) are unable to generate *proof* that an athlete is ‘clean’. Put simply, no test is perfect. The lack of evidence does not mean with absolute certainty that there is no evidence; and equally if evidence is found, it may have a legitimate explanation other than doping. Owing to the potential consequences from a false positive, any testing protocol must err on the side of caution and its diagnostic tests must guarantee a low risk of falsely accusing honest athletes with doping.

Clean Protocol is not an isolated attempt. Developments around anti-doping, which focus on 'the bad athletes', signals a change in directions toward non-analytical forensic approaches -

either in lieu of or to inform the resource-intensive analytical testing. In 2011, the World Anti-Doping Agency bestowed the Young Investigator's Award for the development of the attitude-based “Forensic Anti-Doping Interview, or FADI, as a standardised diagnostic assessment tool that can be used to identify athletes who may be using banned substances” (James Cook University News & Media, 2010; The Profiler, 2011). Even though implementation has not been attempted, anti-doping organisations have had a natural interest in methods - analytical, forensic or psychological - that are capable of identifying doping users. In the past five years, anti-doping agencies funded research into exploring the usefulness of indirect approaches to detect doping behaviour, such as the false consensus effect (Petróczi, Mazanov, Nepusz, Backhouse & Naughton, 2008; Uvacsek, Ránky, Nepusz, Naughton, Mazanov, & Petróczi, 2011) and the implicit association concept (Brand, Heck & Ziegler, 2014; Brand, Wolff & Thieme, 2014; Petroczi, Aidman & Nepusz, 2008; Petróczi et al, 2011). Whilst these attempts did not fulfil the need of producing a diagnostic tool, the results offered valuable insights into athletes’ doping mindsets and highlighted the complexity that surrounds the detection of doping behaviour with psychometric testing.

Doping research includes exploratory work in personality profiling of doping users and of athletes who are susceptible for doping (e.g., Barkoukis, Lazuras, Tsorbatzoudis, & Rodafinos 2011; Gucciardi, Jalleh, & Donovan, 2011); along with identifying the ingredients of a doper prototype (Whitaker, Long, Petróczi, & Backhouse, 2013), but without validated psychometric measures. Although not involving psychometrics, ‘muscle profiling’ (identifying suspects based on having unusually large muscles) is an accepted practice of some police units and national anti-doping organisations (Mulrooney & van de Ven, 2015).

Taken all together, it is vital that anti-doping organisations engaging in doping control, prevention and/or education willingly providing services in support for pre-emptive actions against doping - and their customers - are cognisant and cautious about the limitations inherent in direct and indirect psychometric measurements. Recent developments in anti-doping - with informed, intelligence-led approach and targeted testing within the anti-doping programme - further underscore the need for a global guidance on psychometric testing.

In this commentary, as a group of leading European experts in psychological research of doping in sport, we present a position statement calling for policy guidance on appropriate use of psychometric assessments in anti-doping. We argue that (1) these measurements have been designed, tested and validated to explore how groups of athletes *feel* and *think* about doping, not to determine whether an individual athlete engages in prohibited performance enhancing practices, (2) the psychometric properties established for the athletic groups in controlled research settings under anonymous conditions should not be interpreted as ecological validity for individual diagnostics and (3) the unique characteristics of athletic populations at different levels of involvement and doping-control (e.g., elite, sub-elite, amateur competitive and recreational) must be taken into consideration when interpreting psychometric test differences. It is imperative that we look at these psychometric tests with critical eyes and set clear boundaries for what each can and cannot be used for. In order to inform and protect athletes, anti-doping officials and policy makers from the consequences of potential misuse of the existing psychometric tests, we provide a succinct critical evaluation of the direct and indirect methodologies used in the context of doping prevention. We then make recommendations for the key ingredients of a global policy guide on the use of psychometric testing in social science doping research and anti-doping.

Psychometric and psychological testing in doping research

The use of psychometric testing in doping research has been limited to testing hypotheses of assumed relations and interactions of cognitive and affective variables, and their sole or synergistic effect on behavioural intention and implementation. The primary aim of this research is to identify social cognitive variables or parsimonious models that best describe an athlete doping mindset (Petróczi, 2013a). This work is still in its infancy with the main focus on the development and validation of direct and indirect psychometric measures that sufficiently operationalise and quantify the most promising structures. To date, doping specific psychometric testing has mainly focused on operationalising and quantifying attitudes toward doping (e.g., Brand, Melzer, & Hagemann 2011; Brand, Heck & et al, 2014; Petróczi & Aidman, 2009; Petroczi, Aidman, & Nepusz, 2008; Petróczi, 2013a). At this stage, none of the existing measures are without limitations, which in turn prevents forensic diagnostic application or profiling. In most cases, applications of these psychometric assessments have been limited to evidencing relative differences (e.g., doping users score higher or respond faster than non-users). The two attempts for establishing cut-off values for identifying doping users are not without limitations either. The combination of explicit attitude and projected doping prevalence is based on self-reported behavioural index (Uvacsek et al, 2011), which limits the validity of the model for those who are willing to admit doping. The cut-off value in the other, response-time based attitude measure, was established on analytical results (Brand, Wolff, et al, 2014), but it lacks specificity (i.e., produces high proportion of false positives). These limitations are discussed later.

Other, previously validated but not doping-specific psychological tests used in doping research centre on the influence of socially desirable responding and sport-specific personality traits, beliefs and motivations. Examples of validated psychometric assessments used in doping research since 2000 are summarised in Table 1. Fitness and recreational sport settings (including

bodybuilding) are excluded because doping from the regulatory point of view has had little relevance outside competitive sport settings and WADA governance. Nonetheless, there is evidence that this aspect may change in the future. In some European Union countries (e.g., Belgium, Denmark, Sweden), anti-doping and general drug control measures are getting closer aligned, with anti-doping efforts - including doping testing - being extended to recreational/fitness gyms (Mulrooney & van de Ven, 2015).

< TABLE 1 IS ABOUT HERE >

It must be noted that the psychometric assessments listed in Table 1 were exclusively used to test hypothesised relationships between social cognitive variables and doping or examine differences between doping users vs. non-users to identify protective and motivating factors. For the overview of other personality and social cognitive variables measured and tested in relation to doping, interested readers should consult the meta-analysis by Ntoumanis, Ng, Barkoukis and Backhouse (2014).

Identifying doping users

Establishing evidence for the presence or absence of doping carries a considerable amount of responsibility. Anti-doping organisations with sanctioning power take a cautious and conservative approach to drug testing because the impact of a false positive result is potentially career- (if not life-) changing. In doping detection, accuracy is a conservative balance between sensitivity and specificity that favours the latter. Social, ethical and legal ramifications of false positives on athletes are far greater than the consequences of a false negative on sport - although one can argue that consequences on clean athletes in the same competition are not to be taken lightly.

Psychology research plays an important role in anti-doping through its contribution to developing better understanding of the reasons and motives behind doping to inform effective and ecologically valid intervention and prevention strategies. However, contribution from psychology research to the array of doping *diagnostic* tools - at least at this point in time - is limited owing to lack of development and validation at the individual diagnostic level. At the same time, there is a growing and thus worrying trend to employ forensic intelligence to doping testing and integrate information from social science, among other sources, into a forensic module of organised intelligence (Marclay, Mangin, Margot & Saugy, 2013).

Beyond the foreseeable consequences of a positive identification as a doping user, this task is further complicated by the discrepancy between what constitutes a doping offence according to the doping control regulation and in social science doping research. From the regulatory point of view, a precisely defined set of substances and methods are deemed to be unacceptable and thus prohibited. Engaging in activities involving these substances and methods equates to a doping offence (World Anti-Doping Code, 2015). From the psychological point of view, the sliding scale of assisted performance enhancement that includes both non-prohibited means, such as nutritional or herbal supplements, and prohibited doping makes psychometric assessment of the performance enhancement and doping related social cognition a delicate task, where framing, phrasing and context can individually and collectively exert significant influence on the test outcomes (Petróczi, 2013a).

Cognitive indicators of doping behaviour

Based on the prevailing assumption that dopers must have a rational doping mindset that leads to and supports doping use, scientific inquiry has generally utilized well-developed theoretical frameworks (e.g., social-cognitive models) to study doping behaviour and its socio-

cognitive determinants and correlates (Johnson, 2012; Ntoumanis et al., 2014). Social scientists have also made attempts to use psychometric measures in lieu of analytical approaches in doping and beyond (Agosta & Sartori, 2013; Uvacsek et al, 2011). Despite the promising preliminary results, a wide range of limitations has been identified outside clinical application (Brand, Wolff, & Thieme, 2014; Petróczi et al., 2011; Takarangi, Strange, Shortland, & James, 2013; Vargo, Petróczi, Shah, & Naughton, 2014) that impedes the use of these methodologies in field settings.

As an overarching issue, first we need to take the tenuous connection between cognitive indicators (e.g., doping attitudes) and doping behaviour into account. Whilst significant relationships between psychological variables and self-reported doping intentions and behaviour have been reported, a recent meta-analysis of the extant literature (Ntoumanis et al., 2014) showed that the effect sizes of these relationships were small (e.g., the effect of attitudes on doping behaviour was 0.17, k=13). Owing to the cross sectional nature of the studies conducted to date, causation cannot be established.

It is also important to note that in investigating the complex relations between social cognitive factors and doping-related behaviours, the social cognitive constructs are latent variables. Because they cannot be directly measured, they are inferred from composite scores of items measuring a psychological construct in explicit methodology or from response latency in implicit tasks. The prediction of doping use and the distinction between users and non-users is then established by statistical equations relying on statistical indices such as means, dispersion and correlation. Importantly, both the measurements and the behavioural models in this scenario contain inherent errors, which are minimised but cannot be entirely eliminated.

Another critically important limitation arising from the characteristic feature of doping behaviour research is that the outcome measure (if an athlete dopes or not) is mostly established

on self-report (Petróczi, 2013a). This feature limits the generalizability of the results to those who are willing to admit, at least under anonymous conditions, that they are involved in prohibited performance enhancement practices. Assuming that the self-reported abstinence group is likely to include at least some athletes who deny their action; responding in a socially desirable way inevitably confounds the non-user group cognitive profile.

From the theoretical point of view, it is conceivable that athletes' behavioural choices about doping, and their thoughts and feelings about this choice, have an imprint on their mental representation of doping, which in turn manifests in both explicit and implicit assessments of attitudes (Petróczi. 2013a). Specifically, these influential and interrelated factors can be grouped as behavioural factors (whether an athlete engages in doping or not) and cognitive factors. Taking a closer look at cognitions, an athlete might come to terms with a decision to dope by legitimising the behaviour (e.g., recovery from injury). In contrast, an athlete might feel that doping, as a condemned behaviour, must be denied under all circumstances. Further complicating matters, the athlete's micro-environment (culture) also exerts an influence on these factors.

Measurements

In the real world, the relationships between social cognitions and behaviour may vary and be influenced by a range of contextual features, such as events, situations, circumstances and individual characteristics. The choice of using doping substances is regulated by a complex system of dynamic relations linking motivations, cognitions, and moral convictions or evaluations (Ntoumanis et al., 2014). The psychometric measures developed to evaluate these constructs become meaningful only if they are considered within the theoretical framework describing the pattern of these dynamic relations.

Researchers' evaluative processes are developed according to a top-down approach by which previous theories and knowledge inform specific assessment algorithms, which are then used to develop questionnaire items. Thus, the possibility that the score from a questionnaire cannot be correctly interpreted independently from the theoretical framework at the basis of its development is given. Furthermore, the complexity of the relations between social cognitive constructs related to doping acquires meaning if one considers the possibility that these relations might be embedded, generated and developed within a system of specific social and interpersonal contexts and situations (Hauw, 2013; Zelli, Mallia, & Lucidi, 2010). The understanding of this complexity might not be well served if adopting a dichotomous perspective by which athletes are categorized as non-substance users or as users, purely and thus inappropriately on the basis of their scores on psychometric measures.

Honesty

The validity of self-reported responses is the joint function of the respondents' willingness and ability to respond honestly. Even if an athlete has reasons and he/she is motivated to answer honestly, it is well known that introspecting, consciously accessing, and accurately reporting thought processes that underlie attitudes, motives and behavioural choices is not an easy task. Athletes, who are under obligation to refrain from prohibited methods when they train and compete, have a compelling reason to provide socially, ethically and legally expected answers about their feelings toward using a prohibited performance enhancing substances or their actions when their athletic career depends on being 'doping clean'. Strategic responding is likely to skew results toward a more expressed negative view of doping than the reality, but in this situation strategic responding alone cannot prove guilt because it can result from pressure and fear of appearing guilty without actually being a doper. In sum, no

psychometric method is immune to manipulation and the options available to mitigate against this (i.e., standard impression scales assessing a person's tendency for impression management and/or socially desirable responding) cannot go beyond flagging that caution is warranted in interpreting the self-reported scores.

Indirect approaches

To overcome socially desirable responding, indirect methods have been introduced into doping research. Whilst incorporating indirect measurement into doping-related social cognition research holds promise, caution is warranted in the interpretation of what these indirect measurements actually capture. Social projections of doping use have evidenced biased perception; where the bias is a function of involvement, sensitivity of the behaviour and the reference frame (i.e., in- and out-group) in which the estimation is solicited (Uvacsek et al., 2011; Petróczi, Mazanov, & Naughton, 2011). Based on the assumption that response-time based implicit measurements are less prone to social desirable responding and are thought to reveal automatic associations or evaluations connected to doping, the Implicit Association Test (IAT) concept has been applied to investigating doping attitudes. To date, these tests showed generally unfavourable implicit attitude toward doping regardless of involvement, IAT based testing was so far unable to add substantial explanation of variance to information that could be assessed with questionnaires (Petróczi, 2013b). One key limitation of this line of investigation is that with a few exceptions (e.g., Petróczi et al., 2011; Brand, Wolff, et al., 2014) the outcome behavioural measure was indexed on self-admission.

Test users should dispel the common misconception that implicit measures are panaceas for avoiding strategic responding, or capturing athletes' true feelings outside awareness and conscious control, to evidence doping behaviour *per se*. Rather, implicit measures are better

conceptualised as reflections of athletes' and the various non-athletic population groups' momentarily captured *thoughts about* doping behaviour. Although the lie detector variants such as the autobiographical IAT (Agosta & Sartori, 2013) and the Timed Antagonistic Response Alethiometer (Gregg, 2007) target specific life events (e.g., doping use), they are not free from a host of contextual factors that can influence test performance and produce false outcomes (Vargo et al., 2014). Finally, implicit tests are not immune to manipulation either. With training in deception, one can fake the reaction-time based test - although doing so is less straightforward than manipulating questionnaire responses. More importantly, research outside doping (Hu, Chen, & Fu, 2012; Takarangi et al., 2013) suggests the possibility of a prolonged period of denial of a socially unaccepted and punishable behaviour affecting implicit associations and mental representations, and thus can genuinely produce 'false' memories. At the current stage of our understanding of the processes that underlie indirect task performance, response-time based indirect tasks are not yet suitable for individual diagnostics but they are useful measures for group-level assessment, particularly in combination with direct measures (Payne & Gawronski, 2010; Petróczi et al., 2011; Petróczi, 2013b; Perugini, Richetin & Zogmaister, 2010; Reinecke, Becker & Rinck, 2010).

A doping specific consideration for lie-detection is the complexity of doping, and its effect on the athlete's doping mindset. The way athletes *think* about doping, and their cognitive consistency between feeling, thinking and doing (or not doing), has a profound effect on how they answer statements of the direct psychometric scale items and how they perform on response time- and/or physiological response-based tests. As an illustration, Figure 1 depicts a simple scenario where the way the doping statement is phrased (which the tested athlete is asked to declare as true or false, to make an estimation or to express a degree of agreement) can produce

different and in some cases not interpretable test outcomes. This influence is expected to be greater for a definite behaviour (e.g., using doping) than it is for thoughts and feelings (e.g., attitudes) or projecting similarities or differences in behaviour; and it is assumed to be less controllable or predictable in reaction-time based 'implicit' assessments than it is in self-reports.

< FIGURE 1 IS ABOUT HERE >

Lack of established norms

In the absence of established generalised reference values for each measurement that separates dopers from non-dopers with acceptable accuracy across the full spectrum of the target population and the full range of scores, individual diagnostics with these psychometric tests at this point is impossible. With two exceptions (Brand, Wolff, et al, 2014; Uvacsek et al, 2011), psychometric assessments related to doping, so far, have not established any cut-off or threshold criteria that could distinguish dopers from non-dopers. Even when some threshold value is set for separating doping users from the clean athletes, these values serve as guidance for future research studies at group level assessments, not for forensic diagnostics of individual athletes. Studies comparing self-admitted doping users to non-users on doping attitudes, intentions to use doping and perceived prevalence consistently showed statistically significant difference (e.g., Barkoukis, Lazuras, Tsorbatzoudis & Rodafinos, 2013; Morente-Sánchez, Femia-Marzo, & Zabala, 2014; Petróczi & Aidman, 2009; Uvacsek et al., 2011; Whitaker, Long, Petróczi, & Backhouse, 2013) with doping users scoring higher than non-users, where high scores indicate stronger intention, more positive attitude and higher perceived doping prevalence. Apart from a potential false telling mechanism (Petróczi & Haugen, 2012), scores on the high end of the respective scale's spectrum are generally accepted as unbiased accounts - which cannot be said about the extremely low scores with the same confidence (Petróczi et al., 2011). However, even

if we assume that athletes respond honestly and self-report questionnaires are unbiased (which conditions are most likely not fully met in making socially and legally sensitive self-reports), it is still impossible to identify a threshold for distinguishing users and non-users based on their self-reported beliefs, cognition or affects. The research outcomes available in the literature are relative results (i.e., always presented in the context of non-users' scores), not context-free, absolute criteria for indicating doping use or classifying athletes as users or non-users based on the scoring.

This is also true for measurements that rely on response time differences in implicit tasks. Comparative studies established relatively stronger preference for doping (compared to supplements) or automatic associations between doping and a positive valence category when contrasted to negative valence (Brand, Heck, & Ziegler, 2014; Brand, Melzer, 2011, & Hagemann; Petróczi, Aidman, & Nepusz, 2008) but similarly to the self-reports, no clear threshold could be established for identifying doping users or predict doping use with sufficient confidence. In a study with bodybuilders that combined pictorial brief Implicit Association Test (BIAT) with urinalyses to confirm doping use or absence, a cut-off value for BIAT score was set to identify dopers (Brand, Wolff, et al., 2014). However, the price for high sensitivity to achieve sufficient power to identify true positive cases (i.e., dopers) was the low specificity, which inevitably yielded false positives, the pictorial doping BIAT and the established cut-off value lack the robustness that is needed for forensic applications.

Attempts outside doping research to use the implicit association concept to identify concealed life events - referred to as autobiographical IAT (aIAT) in which the direction of the score (positive or negative) by definition should be indicative for guilty vs. non-guilty - has failed in field settings (Takarangi et al., 201; Vargo et al., 2014). In a small study (Petróczi,

2013c) of 14 male football players who reported no doping use, the morally framed doping aIAT tests ("I cheated with doping") identified 8/14 players as doping user, of which 5 also had positive D-scores and thus identified as user on the aIAT tests using non-judgemental functional frame ("I enhanced my athletic performance"). One potential explanation is that players did not reveal the truth about their doping. The other - and more likely - explanation is that in the true absence of the target behaviour (e.g., doping), the aIAT does not accurately detect absence but rather, it measures some related concept. Such phenomenon has been documented in previous studies using aIAT to identify cocaine users (Vargo & Petróczi, 2013; Vargo et al., 2014), and indicated by the relatively high rate of false positives (5/14) with the functional-frame aIAT (Petróczi, 2013c). Considering the potentially strong confounding effect from vicarious experiences and other general associations, along with the expected framing effect, re-examination of the lie-detector variants of the implicit tests is advised before the wider application of such instrument is made to identify doping users.

Unclear definitions and the danger of naming fallacy

In order to understand the link between the social cognitive constructs and personality traits assessed by psychometric testing and the actual doping behaviour, we need to be precise and specific about the measured construct; and how it is expected to link with behaviour and other constructs. As a minimum, evidence should be offered for any psychometric scale that it actually measures what it claims to measure. As Table 1 shows, to date the Performance Enhancement Attitude Scale (Petróczi, 2002; Petróczi & Aidman, 2009) is the only established doping-specific self-reported measure for athletes, which continues to accumulate evidence for their validity, reliability and generalizability to become a standard research tool for explicit general doping attitude measure. Sullivan and colleagues (2015) recently proposed the Doping

Confrontation Efficacy Scale, which measures a specific domain of coaching efficacy in confronting doping athletes; and offered preliminary evidence for the scale's reliability and predictive validity.

The pool of indirect assessment, on the other hand, is quite murky. Projected use of doping, despite the clear indication of the egocentric bias, is often taken at face value and interpreted as a more 'honest' estimation of doping prevalence when in fact it is not. Projection tells us more about the person who makes the projection than the phenomenon the projection was solicited for. To navigate in the burgeoning field of the reaction-time based assessments is even more difficult. The relationship between response-time based measures and behaviour is not as well understood as the link between explicitly expressed thoughts and feelings and reported behaviour. Indirect methodology, including lie detectors, should not be used in forensic diagnostic settings until we develop a full understanding of the factors that can influence the individual's results.

Need for a clear distinction between research and practical applications

In doping research, the established explicit and implicit measures are used to understand human behaviour in a bid to inform policy and practice and to explore the cognitive processes that underlie doping-related decisions; whereas practice is mostly concerned with diagnostic power and profiling.

To our best knowledge, the Clean Protocol is the first *institutionalised* attempt for using psychological profiling to identify doping users. Nonetheless, it may not be an isolated, commercially motivated attempt but the start of something potentially problematic.

The seriousness of this problem is evident if athletes - at any point in the future - are required to undergo psychometric testing. Given the sensitivity of the issue and high level of

suspicions of doping in elite sport, participation in such a scheme may not be entirely voluntary and free of pressure or coercion. If this commercial enterprise gains momentum, we fear that athletes might face a situation in which refusing to participate in such a protocol raises suspicion of doping and this has serious implications for the athlete and their support network. We urge anti-doping organisations and policy makers to establish global safeguards through robust policy guidance to prevent ethically questionable practices and to prevent misuse of psychological research tools for diagnosis and/or profiling.

The increased need for outcome based evaluation of anti-doping intervention- and prevention programmes will eventually call for the use of standardised psychometric measures of the targeted psychosocial variables and individual level testing. Potential problems from misusing psychometric testing for individual forensic diagnostics and/or profiling can be prevented or contained if proper safeguards are in place. The proposed policy guidance will foster this important development.

Recommendations for policy guidance on psychological assessments in anti-doping

Following the broad guidance on psychometric testing in research and occupational settings we put forward a list of key principles (Table 2) that collectively should serve, over and above the code of ethics and professional standards, as the cornerstone of a global policy on the use of psychometric testing, or not, as part of anti-doping.

< TABLE 2 IS ABOUT HERE >

In addition, guided by the practice followed by the British Psychological Society for psychometric and psychological testing in occupational, educational and forensic contexts, we recommend that practitioners should obtain qualification in psychometric testing and are

registered as approved practitioners. A list of approved psychometric and psychological tests for forensic diagnostics and profiling in relation to doping behaviour should be kept and clearly distinguished from the validated psychometric instruments intended for research only.

The World Anti-Doping Agency (WADA) with its mission to harmonise global anti-doping activities is well positioned to undertake this task. Perhaps modelled after the established procedure for the inclusion/exclusion of substances into the WADA List of Prohibited substances, the Agency could establish and maintain a compilation of validated psychometric tests (and their adaptations to different cultures and languages) with relevance to doping and anti-doping. This authoritative reference material should contain descriptions and impartial critical evaluations of the validated tests, with clear guide for the intended use, populations for which the test has been validated, population norms (if established) and limitations. An expert group - linked to the Education and the Health, Medical and Research committees - could be established with the function of setting and revising the guidelines for psychological and psychometric testing in anti-doping practice and research; and managing test reviews for inclusion in the proposed list. These standards, set and revised in consultation with all stakeholders annually by the committee, could offer clear guidance for researchers and practitioners involved in anti-doping prevention and research.

Conclusion

A false ‘positive’ psychological profile for doping (or even failing to produce a definite negative profile) affects not only the individual athlete but also their entourage, their organisation and sport itself. With these points in mind, future research should carefully consider whether test validation data is sufficient of the test to be used with the target population. A critical examination of the lie-detector methodology in doping context before its application outside

research settings is also warranted. Both will require a full exploration of the contextual factors that may influence the measurement outcomes and/or produce false positives.

At the moment, no doping specific test exists that is psychometrically sound and valid to detect or predict individual behaviour. The potential framing effect presents challenges to adapting general tests (including lie detectors) to the doping context. Therefore the existing doping specific psychometric and psychological tests or ad hoc adaptation of existing general psychometric and psychological tests to doping should not be used in practice for profiling or forensic diagnostics for doping without appropriate validation for such use. Economically and/or politically motivated stakeholders must be aware of the limitations of the existing tests and avoid over-interpretation of what these tests are capable of. To date, the existing psychometric tests serve well as research instruments but diagnostic tools with respective properties have not yet been established. The Policy guidance is needed to control the use of these measures in doping settings at the individual level and to inform and guide future efforts toward validating psychometric and psychological tools for profiling and forensic diagnostics.

The overarching aim of the proposed policy guidance is to protect the global athletic community against social, ethical and legal consequences from potential misuse of psychological tests, including applications as forensic diagnostic tools in both practice and research. It is imperative that users of these psychological tests - researchers and anti-doping personnel alike - are aware of and respect their limitations. Doping-related psychometric tests measure the outcomes of athletes' thinking processes about doping, not the presence or absence of doping.

References

Agosta, S., & Sartori, G. (2013). The autobiographical IAT: a review. *Frontiers in Psychology, 4*, 519. doi:10.3389/fpsyg.2013.00519.

Allen, J., Taylor, J., Dimeo, P., Dixon, S., & Robinson, L. (2014). Predicting elite Scottish athletes' attitudes towards doping: examining the contribution of achievement goals and motivational climate. *Journal of Sports Sciences, 33*, 899-906. doi: 10.1080/02640414.2014.976588.

Backhouse, S.H., Patterson, L. & McKenna, J. (2012). Achieving the Olympic ideal: Preventing doping in sport. *Performance Enhancement and Health*, 1(2), 83-85. doi:10.1016/j.peh.2012.08.001.

Barkoukis, V., Lazuras, L., Tsorbatzoudis, H., & Rodafinos, A. (2013). Motivational and social cognitive predictors of doping intentions in elite sport: An integrated approach. *Scandinavian Journal of Medicine and Science in Sports, 23*, e330-e340. doi: 10.1111/sms.12068.

Barkoukis, V., Lazuras, L., & Tsorbatzoudis, H. (2014). Beliefs about the causes of success in sports and susceptibility for doping use in adolescent athletes. *Journal of Sports Sciences, 32*, 212-219. doi:10.1080/02640414.2013.819521.

Barkoukis, V., Lazuras, L., Tsorbatzoudis, H., & Rodafinos, A. (2011). Motivational and sportspersonship profiles of elite athletes in relation to doping behavior. *Psychology of Sport and Exercise, 12*, 205-212. doi:10.1016/j.psychsport.2010.10.003.

Bartholomew, K.J., Ntoumanis, N., & Thøgersen-Ntoumani, C. (2010). The controlling interpersonal style in a coaching context: Development and initial validation of a psychometric scale. *Journal of Sport & Exercise Psychology, 32*, 193–216. PMID:20479478.

Beck, A.T. (1967). *Depression: causes and treatments.* Philadelphia, PA: University of Pennsylvania Press.

Boardley, I.D., & Kavussanu, M. (2007). Development and validation of the moral disengagement in sport scale. *Journal of Sport and Exercise Psychology, 29*, 608-628. PMID:18089895.

Boardley, I., & Kavussanu, M. (2008). The Moral Disengagement in Sport Scale - Short. *Journal of Sports Sciences, 26,* 1507–1517. doi:10.1080/02640410802315054.

Brand, R., Heck, P., & Ziegler, M. (2014). Illegal performance enhancing drugs and doping in sport: a picture-based brief implicit association test for measuring athlete's attitudes. *Substance Abuse Treatment, Prevention, and Policy, 9,* 7. doi:10.1186/1747-597X-9-7.

Brand, R., Melzer, M., & Hagemann, N. (2011). Towards an implicit association test (IAT) for measuring doping attitudes in sports. Data-based recommendations developed from two recently published tests. *Psychology of Sport and Exercise*, 12, 250-256. doi:10.1016/j.psychsport.2011.01.002.

Brand, R., Wolff, W., & Thieme, D. (2014). Using response-time latencies to measure athletes' doping attitudes: the brief implicit attitude test identifies substance abuse in bodybuilders. *Substance Abuse, Treatment, Prevention and Policy, 9,* 36, doi:10.1186/1747-597X-9-36.

Cavar, M., Sekulic, D., & Culjak, Z. (2012). Complex interaction of religiousness with other factors in relation to substance use and misuse among female athletes. *Journal of Religion and Health, 51,* 381-389. doi:10.1007/s10943-010-9360-9.

Chan, D. K. C., Dimmock, J. A., Donovan, R. J., Hardcastle, S. A. R. A. H., Lentillon-Kaestner, V., & Hagger, M. S. (2014). Self-determined motivation in sport predicts anti-doping

motivation and intention: A perspective from the trans-contextual model. *Journal of Science and Medicine in Sport, 18*, 315–322. doi: 10.1016/j.jsams.2014.04.001.

Conroy,D.E., Elliot, A.J., & S.M. Hofer, S.M. (2003). A 2 × 2 achievement goals questionnaire for sport. *Journal of Sport and Exercise Psychology, 25*, 456–476. doi: 10.1016/j.psychsport.2005.08.012.

Cornwall, W. (2014). To eliminate dopers: is a lie detector test the answer? 10 October 2014. http://www.outsideonline.com/fitness/bodywork/the-fit-list/To-Catch-a-Doper-Is-a-Lie-Detector-Test-the-Answer.html.

Crowne, D.P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354. doi:10.1037/h0047358.

Duda, J. L., & Nicholls, J. G. (1992). Dimensions of achievement motivation in schoolwork and sport. *Journal of Educational Psychology, 84*, 290–299. doi: 10.1037/0022-0663.84.3.290.

Elbe, A.-M., Schlegel, M. M., & Brand, R. (2012). Psychogenic urine retention during doping controls: Consequences for elite athletes. *Performance Enhancement & Health, 1*, 66-74. doi:10.1016/j.peh.2012.01.001.

Gill, D.L., & Deeter, T.E. (1988). Development of the Sport Orientation Questionnaire. *Research Quarterly for Exercise and Sport, 1988, 59*, 191-202. doi:10.1080/02701367.1988.10605504.

Gregg, A.P. (2007). When vying reveals lying: the Timed Antagonistic Response Alethiometer. *Applied Cognitive Psychology, 21*, 631-647. doi:10.1002/acp.1298.

Gucciardi, D.F., Jalleh, G., & Donovan, R.J. (2011). An examination of the Sport Drug Control Model with elite Australian athletes. *Journal of Science and Medicine in Sport, 14*, 469–476. doi:10.1016/j.jsams.2011.03.009.

Gucciardi, D. F., Jalleh, G., & Donovan, R. J. (2010). Does social desirability influence the relationship between doping attitudes and doping susceptibility in athletes?. *Psychology of Sport and Exercise, 11*, 479-486. doi:10.1016/j.psychsport.2010.06.002.

Hanstad, D. V., & Loland, S. (2009). Elite athletes' duty to provide information on their whereabouts: Justifiable anti-doping work or an indefensible surveillance regime?. *European Journal of Sport Science, 9*(1), 3-10. doi:10.1080/17461390802594219.

Hauw, D. (2013). Toward a situated and dynamic understanding of doping behaviors. In J Tolleneer, S. Sterck & P. Bonte (Eds): *Athletic Enhancement, Human Nature and Ethics* (pp. 219-235). Springer Netherlands. doi:10.1007/978-94-007-5101-9_12.

Hodge, K., Hargreaves, E. A., Gerrard, D., & Lonsdale, C. (2013). Psychological mechanisms underlying doping attitudes in sport: motivation and moral disengagement. *Journal of Sport & Exercise Psychology, 35*(4), 419-432. PMID: 23966451.

Hu, X., Chen,H., & Fu, G. (2012). A repeated lie becomes a truth? The effect of intentional control and training on deception. *Frontiers in Psychology, 3*, 488. doi:10.3389/fpsyg.2012.00488.

James Cook University News & Media (2010). The lies have it: busting drug cheats, December 9, 2010. http://www-public.jcu.edu.au/news/JCUPRD1_069260.

Johnson, M.B. (2012). A systemic social-cognitive perspective on doping. *Psychology of Sport and Exercise, 13*, 317-323. doi:10.1016/j.psychsport.2011.12.007.Laure, P., & Binsinger,

C. (2007). Doping prevalence among preadolescent athletes: a 4-year follow-up. *British Journal of Sports Medicine, 41*(10), 660-663. doi:10.1136/bjsm.2007.035733.

Lonsdale, C., Hodge, K., & Rose, E.A. (2008). The Behavioral Regulation in Sport Questionnaire (BRSQ): Instrument development and initial validity evidence. *Journal of Sport & Exercise Psychology, 30*, 323–355. PMID:18648109.

Marclay, F., Mangin, P., Margot, P. & Saugy, M. (2013). Perspectives for forensic intelligence in anti-doping: thinking outside of the box. *Forensic Science International, 229*, 133-144. doi:10.1016/j.forsciint.2013.04.009.

Morente-Sánchez, J., Femia-Marzo, P., & Zabala, M. (2014). Cross-cultural adaptation and validation of the Spanish version of the Performance Enhancement Attitude Scale (Petróczi, 2002). *Journal of Sports Science & Medicine, 13*, 430. PMID: 24790501.

Mulrooney, K.J.D, & van de Ven, K. (2015). "Muscle Profiling": Anti-doping policy and deviant leisure. Deviant Leisure, January, 2015. https://deviantleisure.wordpress.com/2015/01/21/muscle-profiling-anti-doping-policy-and-deviant-leisure/

Newton, M., Duda, J. L., & Yin, Z. (2000). Examination of the psychometric properties of the perceived motivational climate in sport questionnaire-2 in a sample of female athletes. *Journal of Sports Sciences, 18*, 275–290. doi:10.1080/026404100365018.

Ntoumanis, N., Ng, J. Y., Barkoukis, V., & Backhouse, S.H. (2014). Personal and psychosocial predictors of doping use in physical activity settings: A meta-analysis. *Sports Medicine,* doi:10.1007/s40279-014-0240-4.

Paulhus, D.L. (1988). *Manual for the balanced inventory of desirable responding: version 6*. Unpublished manual, University of British Columbia.

Payne, B.K, & Gawronski, G.B. (2010). A history of implicit social cognition: where is it coming from? Where is it going? In B. Gawronski, & K. Payne (Eds), *Handbook of implicit social cognition: measurement, theory, and applications* (pp. 1-8). New York, NY: Guilford Press.

Pelletier, L., Fortier, M. Vallerand, M., Tuson, K., Briere, N. (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*, 35–53.

Petróczi, A. (2002). *Exploring the doping dilemma in elite sport: Can athletes' attitudes be responsible for doping?*. Published Doctor of Philosophy dissertation, University of Northern Colorado, USA.

Petróczi, A. (2007). Attitudes and doping: a structural equation analysis of the relationship between athletes' attitudes, sport orientation and doping behaviour. *Substance Abuse Treatment, Prevention, and Policy, 2*, 34. doi:10.1186/1747-597X-2-34.

Petróczi, A. (2013a), The doping mindset - Part I: Implications of the functional use theory on mental representations of doping. *Performance Enhancement and Health, 2*, 153-163. doi:10.1016/j.peh.2014.06.001.

Petróczi, A. (2013b). The doping mindset - Part II: Potentials and pitfalls in capturing athletes' doping attitudes with response-time methodology. *Performance Enhancement and Health, 2*, 164-181. doi:10.1016/j.eh.2014.08.003.

Petróczi, A. (2013c). Getting inside the athletes' minds: potentials and pitfalls of self-reports and timed response measures in doping research. Keynote paper presented at the 5th International INHDR Conference, 15-16 August, Aarhus, Denmark. p67.

Petróczi, A., & Aidman, E. (2009). Measuring explicit attitude toward doping: Review of the psychometric properties of the Performance Enhancement Attitude Scale. *Psychology of Sport and Exercise, 10,* 390-396. doi: 10.1016/j.psychsport.2008.11.001.

Petroczi, A., Aidman, E.V., & Nepusz, T. (2008). Capturing doping attitudes by self-report declarations and implicit assessment: a methodology study. *Substance Abuse, Treatment, Prevention and Policy 3,* 9. doi:10.1186/1747-597X-3-9.

Petróczi, A., & Haugen, K.K. (2012). The self-reporting doping game: the paradox of a 'false-telling' mechanism and its implications for research and policy. *Sports Management Review, 15,* 513-517. doi:10.1016/j.smr.2012.04.002.

Petróczi, A., & Nepusz, T. (2011). Methodological considerations regarding response bias effect in substance use research: is correlation between the measured variables sufficient. *Substance Abuse Treatment, Prevention, and Policy, 6,* 1, doi:10.1186/1747-597X-6-1.

Petróczi, A., Mazanov, J., & Naughton, D.P. (2011). Inside athletes' minds: Preliminary results from a pilot study on mental representation of doping and implications for anti-doping. *Substance Abuse, Treatment Prevention and Policy, 6,* 10. doi:10.1186/1747-597X-6-10.

Petróczi, A., Mazanov, J., Nepusz, T., Backhouse, S. H., & Naughton, D. P. (2008). Comfort in big numbers: Does over-estimation of doping prevalence in others indicate self-involvement?. *Journal of Occupational Medicine and Toxicology, 3,* 19. doi:10.1186/1745-6673-3-19.

Petróczi, A., Uvacsek, M., Deshmukh, N., Shah, I., Nepusz, T., Aidman, E.V., Barker, J., Tóth, M., & Naughton, D.P. (2011). Incongruence in doping related attitudes, beliefs and opinions in the context of discordant behavioural data: In which measure do we trust? *PLoS One, 6,* 4. doi:10.1371/journal.pone.0018804.

Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behaviour. In B. Gawronski, & K. Payne (Eds). *Handbook of implicit social cognition: measurement, theory, and applications* (pp. 255-294). New York, NY: Guilford Press.

Pitsch, W. (2013). Tacit premises and assumptions in anti-doping research. *Performance Enhancement & Health, 2*(4), 144-152. doi:10.1016/j.peh.2014.07.001.

Plante, T. G., & Boccaccini, M. T. (1997a). The Santa Clara strength of religious faith questionnaire. *Pastoral Psychology, 45*, 375-387. doi:10.1007/BF02230993.

Plante, T. G., & Boccaccini, M. (1997b). Reliability and validity of the Santa Clara strength of religious faith questionnaire. *Pastoral Psychology, 45*, 429-437. doi:10.1007/BF02310643.

Plante, T. G., Vallaeys, C. L., Sherman, A. C., & Wallston, K. A. (2002). The development of a brief version of the Santa Clara Strength of Religious Faith Questionnaire. *Pastoral Psychology, 50*, 359-368. doi:10.1023/A:1014413720710.

Reinecke, A. Eni S. Becker, E.S., & Rinck, M. (2010). Three indirect tasks assessing implicit threat associations and behavioral response tendencies. Test-retest reliability and validity. *Journal of Psychology, 218*, 4-11. doi:10.1027/0044-3409/a000002.Roberts, G. C., Treasure, D. C., & Balague, G. (1998). Achievement goals in sport: The development and validation of the Perception of Success Questionnaire. *Journal of Sports Sciences, 16*(4), 337-347. doi:10.1080/02640419808559362.

Rosenberg, M. (1979). *Conceiving the self.* New York: Basic Books.

Russell, D., Peplau, L.A., & Cutrona, C.E. (1980). The revised UCLA Loneliness Scale: concurrent and discriminant validity evidence. *Journal of Personality and Social Psychology, 39*, 472-480. PMID:7431205.

Sas-Nowosielski, K., & Swiatkowska, L. (2008). Goal orientations and attitudes toward doping. *International Journal of Sports Medicine, 29*, 607-612. doi:10.1055/s-2007-965817.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Storch, E.A., Kovacs, A.H., Roberti, J.W., Bailey, L.M., Bravata, E.A., & Storch, J.B. (2004). Strength of religious faith and psychological adjustment in intercollegiate athletes. *Psychological Reports, 94*, 48-50. doi:10.2466/pr0.94.1.48-50.

Stöeber, J. (2001). The Social Desirability Scale-17 (SDS-17): convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment, 17*, 222–232. doi:10.1027//1015-5759.17.3.222.

Strahan, R., & Gerbasi, K.C. (1972). Short, homogeneous versions of the Marlowe–Crowne social desirability scale. *Journal of Clinical Psychology, 28*, 191–193. doi:10.1002/1097-4679(197204)28:2<191::AID-JCLP2270280220>3.0.CO;2-G.

Sullivan, P. J., Feltz, D. L., LaForge-MacKenzie, K., & Hwang, S. (2015). The preliminary development and validation of the Doping Confrontation Efficacy Scale. *Psychology of Sport and Exercise, 16*, 182-190. doi:10.1016/j.psychsport.2014.04.011.

Takarangi, M.K.T., Strange, D., Shortland, A.E., & James, H.E. (2013). Source confusion influences the effectiveness of the autobiographical IAT. *Psychonomic Bulletin & Review, 20*, 1232-1238. doi:10.3758/s13423-013-0430-3.

The Profiler (2011). Newsletter of the College of Forensic Psychologists. 2011 April, Volume 11. http://groups.psychology.org.au/Assets/Files/Profiler%20April%2011.pdf

Uvacsek, M., Ránky, M., Nepusz, T., Naughton, D.P., Mazanov, J., & Petroczi, A. (2011). Self-admitted behaviour and perceived use of performance enhancing versus psychoactive

drugs among competitive athletes. *Scandinavian Journal of Science and Medicine in Sport 21*(2), 224-234. doi:10.1111/j.1600-0838.2009.01041.x.

Vallerand, R.J., Briere, N.M., Blanchard, C., & Provencher, P. (1997). Development and validation of the multidimensional sportspersonship orientations scale. *Journal of Sport and Exercise Psychology, 19*, 197–206.

Vargo, E.J., & Petroczi, A. (2013). Detecting cocaine use? The autobiographical Implicit Association Test (aIAT) produces false positives in a real-world setting. *Substance Abuse, Treatment, Prevention and Policy, 8,* 22. doi:10.1186/1747-597X-8-22.

Vargo, E.J., Petróczi, A., Shah, I., & Naughton, D.P. (2014). It is not just memory: Propositional thinking influences performance on the autobiographical Implicit Association Test. *Drug and Alcohol Dependency, 145,* 150-155. doi:10.1016/j.drugalcdep.2014.10.008.

Whitaker, L., Long, J., Petróczi, A., & Backhouse, S. H. (2013). Using the prototype willingness model to predict doping in sport. *Scandinavian Journal of Medicine & Science in Sports*, 24, e398-e405. doi:10.1111/sms.12148.

World Anti-Doping Code. January 2015. https://wada-main-prod.s3.amazonaws.com/resources/files/wada-2015-world-anti-doping-code.pdf

Zelli, A., Mallia, L., & Lucidi, F. (2010). The contribution of interpersonal appraisals to a social-cognitive analysis of adolescents' doping use. *Psychology of Sport and Exerc*ise, *11*, 304-311. doi:10.1016/j.psychsport.2010.02.008.

Zenic, N., Stipic, M., & Sekulic, D. (2013). Religiousness as a factor of hesitation against doping behavior in college-age athletes. *Journal of Religion and Health, 52*, 386-396. doi:10.1007/s10943-011-9480-x.

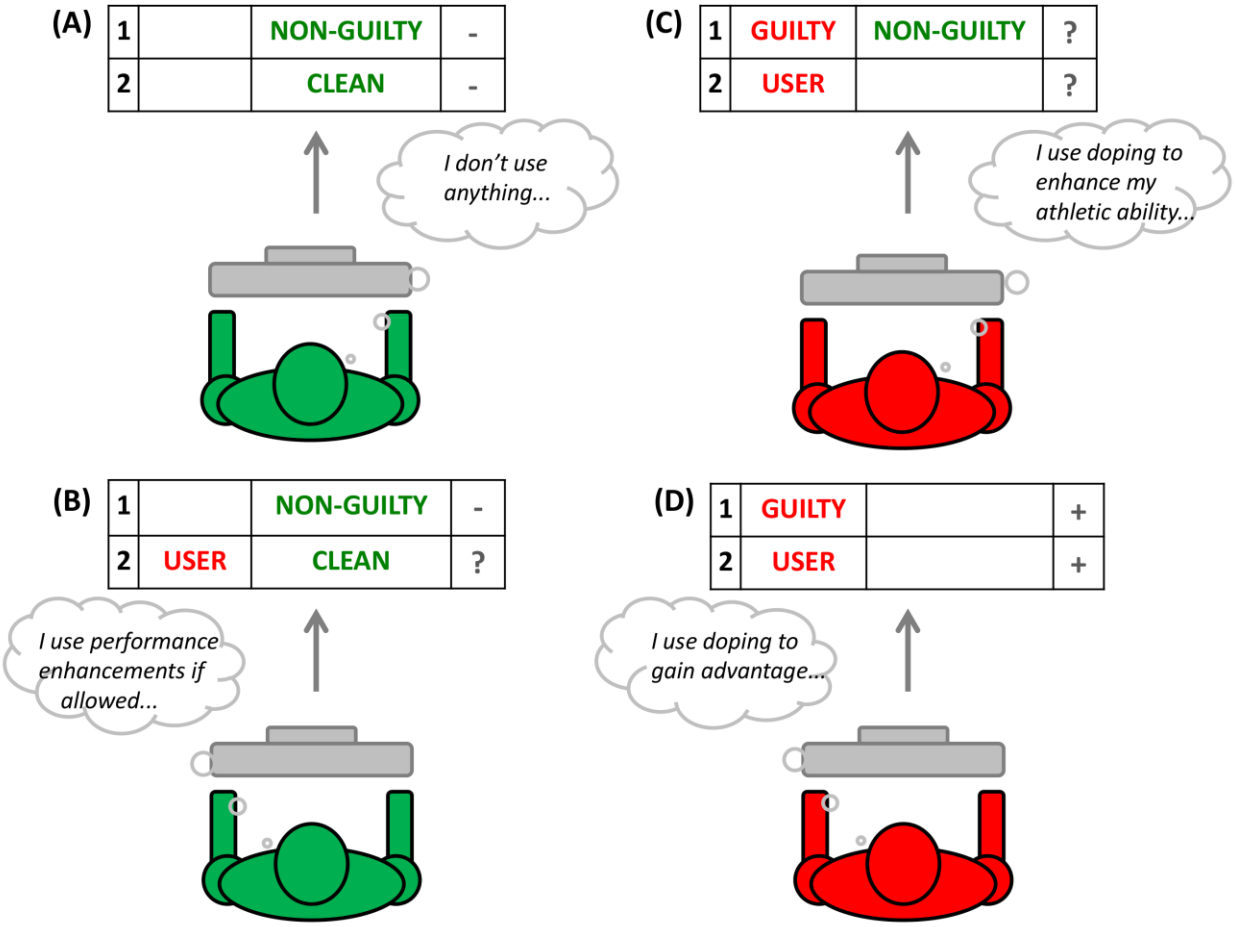

Figure 1: Hypothetical scenario of athletes with different mental representation about doping undergoing psychological assessment. Green figures (A & B) represent clean athletes**;** red figures (C & D) represent doping users. The doping mental representation of athlete C is centred on performance enhancement representing functional thinking and internal motivation, athlete D's doping mindset focuses on gaining advantage in competitive situation, is set in moral framework and representing external motivation. Boxes represent the test outcomes, where 1 is a test with life event statements are framed in judgemental terms (e.g., *"I cheated by using doping"*), 2 is a test where life event statements are factual and non-judgemental (e.g., *"I used prohibited performance enhancing substances to enhance my performance"*); signs represent doping attitudes ([+] tolerant/permissive; [-] intolerant/prohibitive; [?] ambivalent). The figure shows that incongruence between the test frame and the individual's mental representation can produce misleading or ambivalent results.

Table 1

Psychometric assessments used in doping research between 2000 and 2015

	Measure	**Description**	**Reference**[a]
Doping Explicit	Performance Enhancement Attitude Scale (PEAS)	Validated self-reported 17-item explicit measure of general doping attitude	Petróczi, 2002; Petróczi & Aidman, 2009
	Doping Confrontation Efficacy Scale (DCES)	Validated 21-item self-reported measure of a doping-specific domain of coaching efficacy	Sullivan, Feltz, LaForge-MacKenzie, & Hwang, 2015
Doping Implicit[b]	Doping Implicit Association Test	Measure of attribute association valence (affective implicit doping attitude); some evidence for validity	Petróczi, Aidman, & Nepusz, 2008
	Doping Implicit Association Test	Measure of relational target association (affective implicit doping attitude); some evidence for validity	Brand, Melzer, & Hagemann 2011
	Doping Brief Implicit Association Test	Measure of relational target association (affective implicit doping attitude); some evidence for validity in athletes who admit doping	Petroczi et al, 2011
	Pictorial Doping Brief Implicit Association Test	Measure of attribute association valence (affective implicit doping attitude); some evidence for diagnostic power	Brand, Heck & Ziegler, 2014 Brand, Wolff, & Thieme, 2014
Social desirability	Balanced Inventory of Desirable Responding - Impression Management subscale (BIDR-IM)	Validated self-reported measure of social desirability, 20-item Impression Management dimension (Paulhus, 1988)	Whitaker, Long, Petróczi, & Backhouse, 2013
	Marlowe-Crowe Social Desirability Scale (M-C SDS)	Validated 33-item self-report measure of the need to respond in culturally accepted ways (Crowne & Marlowe, 1960)	Petróczi & Nepusz, 2011
	Marlowe–Crowne Social Desirability Scale (SDS); Short version	Validated 10-item version of M-C SDS (Strahan & Gerbasi, 1972)	Barkoukis, Lazuras, Tsorbatzoudis, & Rodafinos, 2011 Barkoukis, Lazuras, Tsorbatzoudis,, & Rodafinos 2013
	The Social Desirability Scale-17	Validated 17-item measure of socially desirable	Gucciardi, Jalleh, & Donovan,

	(SDS-17)	responding (Stöeber, 2001)	2010
General	Approach and Avoidance Achievement Goal Questionnaire (AAAGQ)	Validated 4-dimensional 12-item self-reported measure of mastery-approach/avoidance, performance-approach/avoidance goals (Conroy, Elliot, & Hofer, 2003)	Barkoukis et al, 2011
	Beck Depression Inventory	Validated 21-item self-reported measure of depression (Beck, 1967)	Storch, Kovacs, Roberti, Bailey, Bravata, & Storch, 2004
	Rosenberg's Self-esteem Scale	Validated 10-item self-reported measure of global self-esteem (Rosenberg, 1979)	Laure & Binsinger, 2007 Morente-Sánchez, Femia-Marzo, & Zabala, 2014
	Santa Clara Strength of Religious Faith Questionnaire (SCSORF) - brief version	Validated 5-item brief version of the SCSORF (Plante, Vallaeys, Sherman, & Wallston, 2002).	Storch, Kovacs, Roberti, Bailey, Bravata, & Storch, 2004
	State-Trait Anxiety Inventory (Trait Anxiety subscale)	Validated self-reported measure of state-trait anxiety, 20-item subscale measures stable trait anxiety (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983)	Laure & Binsinger, 2007 Storch, Kovacs, Roberti, Bailey, Bravata, & Storch, 2004
	UCLA Loneliness Scale	Validated 20-item self-reported measure of loneliness (Russel, Peplau, & Cutrona, 1980)	Storch, Kovacs, Roberti, Bailey, Bravata, & Storch, 2004
Sport	Behavioral Regulation in Sport Questionnaire-6 (BRSQ-6)	Validated 24-item self-reported measure of six types of motivational regulation (Lonsdale, Hodge, & Rose, 2008)	Hodge et al, 2013 Chan, Dimmock, Donovan, Hardcastle, Lentillon-Kaestner, & Hagger, 2014
	Beliefs about the Causes of Success in Sport Questionnaire (BACSSQ)	Validated multidimensional 18-item self-reported measure of athletes' perceptions about the causes of success (Duda & Nicholls, 1992)	Barkoukis, Lazuras, & Tsorbatzoudis, 2014
	Coach Controlling Behaviors Scale (CCBS)	Validated, 15-item self-reported measure of the controlling dimension of coaching style/climate (Bartholomew, Ntoumanis, & Thøgersen-Ntoumanis, 2010)	Hodge et al, 2013
	Moral Disengagement in Sport Scale - Short	Short version of the validated MDSS (Boardley & Kavussanu, 200)	Hodge, Hargreaves, Gerrard, & Lonsdale, 2013
	Moral Disengagement in Sport Scale	Validated multidimensional, sport specific 32-item	

	(MDSS)	self-reported measure of moral disengagement (Boardley & Kavussanu, 2007)	
	Multidimensional Sportspersonship Orientation Scale (MSOS)	Validated multidimensional 25-item self-reported measure of five different types of sportspersonship orientations (Vallerand, Briere, Blanchard, & Provencher, 1997)	Barkoukis et al, 2011
	Perceived Motivational Climate in Sport Questionnaire (PMCSQ-2)	Validated 33-item self-report multi-dimensional Measure of perceived motivational climate (Newton, Duda, & Yin, 2000)	Allen, Taylor, Dimeo, Dixon, & Robinson, 2014
	Perceptions of Success Questionnaire (POSQ)	Validated 12-item self-reported measure of dispositional goal orientation (Roberts, Treasure, & Balague, 1998)	Sas-Nowosielski & Swiatkowska, (2008)
	Santa Clara Strength of Religious Faith Questionnaire (SCSORF)	Validated 10-item measure of the strength of religious faith (Plante, & Boccaccini, 1997a; 1997b)	Cavar, Sekulic, & Culjak, 2012 Zenic, Stipic, & Sekulic, 2013
	Sport Motivation Scale (SMS)	Validated multidimensional 28 item self-reported measure of sport-specific motivation (Pelletier, Fortier, Vallerand, Tuson, & Briere, 1995)	Barkoukis et al, 2011
	Sport Orientation Questionnaire (SOQ)	Validated multidimensional, sport specific, 25-item explicit measure of individual differences in sport achievement orientation as competitiveness, winning, and goals (Gill & Deeter, 1988)	Petróczi, 2007
	Task and Ego Orientation in Sport Questionnaire (TEOSQ)	Validated 13-item 2-dimensional self-reported measure of task- (7 items) and ego- (6 items) related goal orientation (Duda, 1989; Duda & Nicholls, 1992)	Allen, Taylor, Dimeo, Dixon, & Robinson, 2014

Notes:

[a] For non-doping specific measures, reference is given where used in doping research.

[b] Depending on the pairing and whether a target concept or an attribute is set as the non-focal category, the IAT measures different constructs. When two targets are contrasted using the same (usually positive) attribute, the IAT outcome is a "relational target association" (e.g., preference [good] for supplements over doping or vice versa). When two attributes are used in combination with a single target category, the IAT measures the strength of the attribute association valence (e.g., doping is more good than bad, or vice versa).

Table 2

Key principles for using psychometric testing in anti-doping.

KEY PRINCIPLES FOR USING PSYCHOMETRIC TESTING IN ANTI-DOPING[a]
Test selection: Test administrators must be able to justify the selection of test(s) and the test(s) selected should be validated and fit for purpose.
Integrity: Modification of a validated test must be documented. Modified test(s), depending on the extent of the changes, may need to be validated.
Use of information: Test administrators have the ethical obligation to ensure that the information arising from psychometric testing is not misused. Therefore, there must be careful attention paid to the interpretation of the information and in defining its limitations**.**
Transferability and generalizability: Outcomes of psychometric testing should not be interpreted without the context in which the data were collected and care should be taken to not place undue weight on the predictive validity of the findings in a different setting unless generalizability has been already established.
Competence: Test administrators must have the minimum necessary competence to make justified test selection, to conduct psychometric testing and to interpret the results.
Scope: Psychometric tests that are not validated for individual diagnosis should not be used for collecting information on individuals.
Psychological assessment linked to intervention: Tests selected should be fit for the purpose required. So psychometric testing employed to provide support for intervention effects should measure constructs that directly map onto the planned intervention outcomes.
Reporting the results: Test administrators must ensure that the reported results are closely linked to the objectives of the assessment and limitations of the test(s) and its effect on the outcome are clearly recognized and reported.
Decision making criteria: Decision about an individual should not be solely based on the outcome from a psychometric test and alternative explanation(s) and interpretation(s) are considered before a conclusion is made.
Access to psychometric tools: If the psychometric tool is in the public domain and freely available, deliberate manipulation and training must be considered in interpreting the results.
Informed consent: Informed consent should be gained from the individual prior to the test being administered. The consent should define who has the right to receive this information. This consent should be in writing, particularly where the information is to be given to a third party.
Education: Test administrators have ethical and legal obligation to provide accurate and specific information about what the psychometric assessment can and cannot do as part of the consenting process.
Confidential storage of test data: Psychometric information should be stored in a secure and confidential manner. The test data should not be accessible to those who are not trained to interpret and should be viewed only by those who have consent from the individual to access the information.

[a] For general guidance on psychometric testing, readers should refer to the American Psychological Association's testing guidelines [http://www.apa.org/science/programs/testing/index.aspx]