Journal of Sport & Exercise Psychology, 2014, 36, 347-356
http://dx.doi.org/10.1123/jsep.2013-0051

JOURNAL OF
SPORT & EXERCISE
PSYCHOLOGY
Official Journal of NASPSPA
www.JSEP-Journal.com
ORIGINAL RESEARCH

Separating Chokers From Nonchokers: Predicting Real-Life Tennis Performance Under Pressure From Behavioral Tasks That Tap Into Working Memory Functioning

Erik Bijleveld[1,2] and Harm Veling[2]
[1]Utrecht University; [2]Radboud University Nijmegen

To better understand the characteristics of athletes who tend to underperform under pressure, we investigated how (a) working memory (WM) capacity and (b) responsiveness of the dopamine system shape real-life performance under pressure. We expected that athletes with smaller WM capacity or a more responsive dopamine system (as operationalized with a risk-taking measure) are especially prone to fail during decisive moments. In a sample of competitive tennis players, WM capacity was measured with the Automated Operation Span task (AOSPAN); responsiveness of the dopamine system was measured with a risk-taking measure, the Balloon Analogue Risk Task (BART). As expected, higher AOSPAN scores predicted better performance during decisive sets; higher BART scores predicted worse performance during decisive sets. These findings indicate that real-life tennis performance can be predicted from behavioral tasks that tap into WM functioning and risk taking, and suggest that the ability to effectively use WM despite pressure separates chokers from nonchokers.

Keywords: choking under pressure, dopamine, risk taking, decisive moments, AOSPAN, BART

Some moments in sports matches are more important than others, and athletic success is strongly related to people's ability to thrive during decisive moments. For many, however, optimal performance during decisive moments is difficult to achieve. Referred to as *choking under pressure*, people tend to perform worse than expected given their ability when they are highly motivated to succeed (Beilock & Gray, 2007). Intriguingly, anecdotal and psychological evidence indicates that some athletes more than others are likely to fail when it matters most (Beilock, 2010; Gucciardi, Longbottom, Jackson, & Dimmock, 2010; Mesagno & Marchant, 2013; Wang, Marchant, Morris, & Gibbs, 2004). The present research addresses the psychological and biological characteristics of these athletes. Specifically, building on the idea that pressure disrupts working memory (WM) functioning, we test whether competitive tennis players' performance during decisive moments can be predicted by athletes' ability to effectively make use of WM when the pressure is on.

The present research contributes to the literature in two ways. First, previous research has often used questionnaires to examine choking-related individual differences. However, WM processes that cause choking are difficult to capture using this methodology. Instead, the present research uses two well-validated behavioral tasks to zoom in on the processes that shape performance under pressure. Second, in the present research, athletes' performance under pressure is deduced from actual match scores, by contrasting tennis players' performance during decisive versus nondecisive sets. A major advantage of this approach is that performance is measured in its natural context.

Erik Bijleveld and Harm Veling are at the Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands. Address author correspondence to Erik Bijleveld at e.bijleveld@psych.ru.nl.

The Psychological Underpinnings of Choking Under Pressure

Two types of theories, thought to be complementary, are often employed to explain how choking under pressure occurs. First, *conscious processing theory* (Masters, 1992) proposes that pressure causes athletes to deploy attention to step-by-step components of motor skills that are normally executed routinely and outside of awareness. This dysfunctional way of using attention disrupts performance on tasks that rely on well-learned movements (Baumeister, 1984; Beilock & Carr, 2001; Masters, 1992). Second, *distraction theory* proposes that performance pressure prompts task-irrelevant thoughts and worries to occupy WM (Baddeley, 2003; Eysenck & Calvo, 1992). Thus, useful processes (e.g., deciding where to hit a tennis ball) compete for limited resources with nonuseful ones (e.g., thinking about the consequences of losing), thwarting performance on tasks that rely on WM and executive functioning (Beilock, Kulp, Holt, & Carr, 2004; Markman, Maddox, & Worthy, 2006; Sorg & Whitney, 1992).

In line with these ideas, previous research addressing individual differences in choking mainly focused on athletes' tendencies to self-monitor or to get distracted. Using questionnaires, such research pointed to roles for *reinvestment*, *self-consciousness*, and *trait anxiety* (Baumeister, 1984; Jackson, Ashford, & Norsworthy, 2006; Masters, Polman, & Hammond, 1993; Wang et al., 2004). Although this approach has produced important insights, a limitation of self-report questionnaires is that they rely on people's (questionable) ability to accurately reconstruct the operation of psychological and brain processes (Nisbett & Wilson, 1977). Thus, questionnaire scores are rather remote from the executive processes that are the basis of performance under pressure. The present study aims to provide new insights into what separates chokers and nonchokers by addressing more proximal causes of choking. Using well-validated tasks, we aim to quantify athletes' ability to employ WM functions effectively when they are under high pressure to perform.

In sports in which performance is a function of quick and high-quality strategic decisions, such as tennis, performance depends on WM. WM, the system that maintains task-relevant information and performs operations on it (Baddeley, 2003), is involved in many aspects of human functioning (Barrett, Tugade, & Engle, 2004) including strategic aspects of sports (Vestberg, Gustafson, Maurex, Ingvar, & Petrovic, 2012). Under pressure, however, people's WM functions are often found to be disrupted, as task-irrelevant thoughts and worries impose load on WM (Beilock, Kulp, Holt, & Carr, 2004), which may impair the efficiency of WM-dependent strategic decision making (Eysenck & Calvo, 1992). Accordingly, we expect that athletes with greater WM capacity are more resistant to choking under pressure. As they have more capacity to begin with, it may for them be less problematic if some of this capacity is compromised (Hypothesis 1).

Research on the neurobiological mechanisms that underpin WM raises further predictions about athletic performance under pressure. WM is mainly supported by the prefrontal cortex (PFC; Miller & Cohen, 2001), a structure that normally orchestrates and coordinates activity in the rest of the brain (Barrett et al., 2004). Importantly, functioning of the PFC is modulated by the neuromodulator *dopamine*, which is known to be released when people are motivated to perform. Up to a moderate level of release, dopamine enhances the extent to which the PFC exerts control over other brain areas. This enhancement of PFC control increases performance on most tasks and increases alertness more generally. However, the relation between dopamine release and PFC control follows an inverted-U shape: when prefrontal dopamine levels reach a tipping point, the PFC no longer coordinates and controls neural activity—instead, it loses control over other brain structures and people come to rely on low-level, emotional reflexes (e.g., mediated by the amygdala; Arnsten, 2009).[1] When this happens, performance suffers on various tasks that require WM and efficient attention regulation (Cools & Robbins, 2004).[2]

It is important to note that the sensitivity of people's dopamine system differs strongly across individuals (e.g., Buckholtz et al., 2010). That is, some people more than others quickly reach the point at which the PFC loses control (i.e., the prefrontal-dopamine tipping point; Arnsten, 2009). As a result, some people are especially prone to dopamine-related decreases in performance. Extending this line of reasoning to athletes' performance, we propose that athletes who have a more sensitive dopamine system are more likely to choke under pressure (Hypothesis 2).

The idea that high levels of dopamine release affect performance by thwarting WM suggests a further possibility. That is, one could argue that—in athletes with a *more* responsive dopamine system—performance under pressure is *less* dependent on WM capacity. As addressed above, these athletes' dopamine system has a stronger tendency to thwart the PFC, and thus WM, during high-pressure situations (see Hypothesis 2; Arnsten, 2009). Taking this idea one step further, one could propose that, when the stakes are high, these athletes' initial WM capacity is no longer predictive of performance. After all, under pressure, their WM is likely to be thwarted anyway due to their responsive dopamine system—whether their WM capacity was high or low to begin with becomes irrelevant. Based on this reasoning, it can be predicted that individual differences in WM capacity and in dopamine responsiveness predict performance in an interactive way. Specifically, it could be the case that WM capacity is less predictive for performance of people who have a more responsive dopamine system (Hypothesis 3).

The Present Research

To test these hypotheses, we conducted a cross-sectional study using a sample of competitive tennis players. As a measure of WM capacity, the Automated Operation Span (AOSPAN) task was administered (Unsworth, Heitz, Schrock, & Engle, 2005). Modeled after the well-known Operation Span Task, the AOSPAN was designed to measure WM capacity without the continuous involvement of an experimenter. In a validation study (Unsworth et al., 2005), the AOSPAN was shown to have good test–retest reliability (r = .83) and internal consistency (α = .78). Specifically, the AOSPAN measures people's ability to hold information in mind while at the same time performing demanding cognitive operations (in this case, mathematical operations).

As a measure of responsiveness of the dopamine system, the Balloon Analogue Risk Task (BART) was administered (Lejuez et al., 2002). The reason we chose a risk-taking task to tap individual differences in responsiveness of the dopamine system is that risk-taking behavior is a behavioral consequence of dopamine system functioning. There are several previous findings that support this idea. First, people who have a more sensitive dopamine system (as indicated by them having genetic variants that directly affect dopamine reactivity) show greater responses to potential rewards in the ventral striatum (Forbes et al., 2009), which in turn drive risky decisions, at least in healthy people (Knutson, Wimmer, Kuhnen, & Winkielman, 2008). Second, people who have a more sensitive dopamine system (as indicated by them having a greater binding potential for dopamine in subcortical areas, and as indicated by them being more responsive to amphetamine) also tend to have a more impulsive personality (Buckholtz et al., 2010), which is closely related to real-life risky behavior (e.g., Martins, Tavares, da Silva Lobo, Galetti, & Gentil, 2004). Third, when people take dopamine agonists to help them deal with Parkinson's disease, this may induce them to develop pathological gambling tendencies (Driver-Dunckley, Samanta, & Stacy, 2003). In other words, artificially boosting the dopamine system causes risk-taking behavior. These prior findings converge on the idea that sensitivity of the dopamine system is a key biological cause of risk-taking behavior.

The reason we chose the BART, specifically, is that it is a reliable and valid measure of risk-taking behavior. Previous validation studies have shown that it has good test–retest reliability (r = .77; White, Lejuez, & de Wit, 2008) and that BART scores correlate well with both self-report and behavioral measures of risky behavior (Lejuez, Aklin, Zvolensky, & Pedulla, 2003; Lejuez, Aklin, Jones et al., 2003). A further asset of the BART is that it has successfully been used in various populations (Fein & Chang, 2008), even different species (Jentsch, Woods, Groman, & Seu, 2010).

Connecting the lines of reasoning addressed above, we conclude that (a) individual differences in sensitivity of the dopamine system cause parallel individual differences in risk taking behavior and that (b) the BART is a high-quality measure of risk-taking behavior. By deduction, BART scores can be assumed to indirectly reflect individual differences in sensitivity of the dopamine system.

In further support of this assumption, several studies directly connect BART scores to functioning of the dopamine system. First, in an fMRI study, risky choices in the BART were associated with activity in the main targets of the mesolimbic and mesocortical dopamine pathways (e.g., the ventral striatum; Rao, Korczykowski, Pluta, Hoang, & Detre, 2008). Second, another study that used the BART suggests that people who have a sensitive dopamine system (as indicated by them having a greater binding potential for dopamine in subcortical areas) show reduced decision-related activity in the dorsolateral prefrontal cortex (Kohno et al., in press), a brain structure known to be involved in suppressing people's inclinations to take risks during the BART (Fecteau et al., 2007). So, although more research is still needed, existing research strongly points to the idea that sensitivity of the dopamine system can be captured with the BART. For that reason, the BART is in our view a promising candidate to predict real-life dopamine-related behavioral outcomes, such as athletic performance.

Athletes' performance under pressure was deduced from tennis set scores from competitive matches. As choking is defined as lower performance in high-pressure compared with low-pressure situations (within-athletes; Beilock & Gray, 2007), we computed a *choking index* that reflected relative performance during decisive sets (i.e., sets in which the match could be ended) compared with nondecisive sets. To test our hypotheses, we regressed tennis players' choking index on their AOSPAN and BART scores.

Method

Sample

Forty-five competitive tennis players took part in the study. They were recruited via targeted advertisements in tennis clubs and via word of mouth. Due to a software error, data from the first seven participants were not usable. For two further participants, scores from only few matches (<6) were available; these participants were a priori excluded from further analysis. The remaining 36 participants (mean age = 24.9, 11 women) had on average 16.1 years of experience playing tennis. Their mean playing level for singles, expressed as the International Tennis Number (ITN), was 3.3 (SD = 1.7). The ITN system is an international system that is used to quantify any given tennis player's level on a scale from 1 (internationally ranked professional) to 10 (beginner). As such, the present sample can be characterized as consisting of "advanced players." All participants competed at the regional or national level. They reported that tennis was an important activity to them (M = 4.2 on a 5-point scale), and that they felt it was important to win (M = 4.3). Participants received €7 for their participation.

Procedure

The two tasks (see below) were administered at the athletes' homes, via a notebook computer. To make sure that participants could focus on the tasks without distraction, the computer was placed against a cardboard screen. Furthermore, participants wore soundproof headphones throughout the testing session, which took around 45 min. The study was conducted according to institutional guidelines and approved by the local ethics committee. All participants gave written informed consent, and permitted us access to their tennis scores. Then, the AOSPAN and the BART were administered.

Automated Operation Span Task (AOSPAN)

The AOSPAN is a computerized version of the often-used Operation Span Task that measures WM capacity.

Participants first saw an arithmetic problem that they had to solve (e.g., "(4 × 3) + 4 = ?"). Participants were instructed to solve the problem as quickly as possible, and to click the mouse when they were ready. Then, participants were shown a number (e.g., "16"), of which they had to indicate whether it was the correct solution to the arithmetic problem. Next, they saw a letter for 800 ms, which they had to maintain in memory. This sequence (solve arithmetic problem, maintain letter in memory) was presented three, four, five, six, or seven times before participants were prompted to report the three to seven letters that they had seen until then. After reporting the letters, participants received feedback on their performance (how many arithmetic problems were solved correctly, how many letters were remembered correctly). In total, the task consisted of 75 arithmetic problems and letters, subdivided into three series of each length (i.e., 3 × 3, 3 × 4, 3 × 5, 3 × 6, 3 × 7).

The AOSPAN score, then, was computed as the total number of letters in series that were remembered correctly (Unsworth et al., 2005). For example, if participants correctly remembered a series of five letters, this was scored as 5 points. Series in which one or more errors were made were scored as 0 points (see e.g., Ahmed & de Fockert, 2012; Decaro, Thomas, & Beilock, 2008, for this scoring procedure). Performance on the arithmetic problems was not included in the score, but participants were instructed to be at least 85% accurate on the arithmetic problems. So, the AOSPAN measures peoples capability of maintaining information in WM, while they at the same time engage in distracting activities. For more details about this task (e.g., about timing and about practice blocks), we refer the reader to Unsworth et al. (2005).

Balloon Analogue Risk Task (BART)

The BART is used to measure people's tendency to take risks—a tendency that is known to mirror sensitivity of the dopaminergic system. In this task, participants were presented with a virtual balloon on the computer screen. Participants could choose to blow air in this balloon by pressing the spacebar, increasing the balloon's size. Every time they chose to do so, this yielded them €0.05. This way, they could accumulate money. By blowing air in the balloon, however, they also ran the risk of popping it. Specifically, with the first blow of air, the chance of popping the balloon was 1/128; with the second, 1/127; with the third, 1/126; and so on until the 128th blow, with which the chance was 1/1. If the balloon popped, all the money that had accumulated so far was lost. So, participants continuously faced the choice between a risky option (blowing air in the balloon) and a safe option (keeping the money they accumulated and starting over with a new balloon).

Participants were notified in advance that they would get 30 balloons to earn money. Although they did not receive the money they accumulated, the participant with the best score received a gift card worth €50 in addition to their regular payment. The BART score was computed by averaging the number of times participants pumped per balloon, across all balloons that did *not* pop (Lejuez, Aklin, Jones, et al., 2003). As such, BART scores linearly reflected people's tendency to take risks, and thus, as we argued above, the sensitivity of their dopamine system.

The reason for excluding scores from balloons that popped was methodological (Lejuez et al., 2002). Specifically, when a given balloon pops, it is no longer possible for participants to choose to pump more air into that balloon—even though they might have done so had the balloon stayed intact. So, participants' scores from these balloons are necessarily constrained, and therefore do not necessarily reflect people's real tendency to take risks. For that reason, scores from popped balloons were a priori excluded from further analyses (see Lejuez et al., 2002; Lejuez, Aklin, Jones, et al., 2003). On average, people let 10.0 out of 30 balloons pop ($SD = 3.7$). For more details about this task, we refer the reader to Lejuez, Aklin, Jones, et al. (2003).

Data Reduction of Tennis Set Scores

We were able to collect set scores for 691 tennis matches played by the 36 players in our sample ($M = 19.2$ matches per participant, $SD = 7.3$). These data were extracted from the Royal Dutch Lawn Tennis Association's database. To interpret these set scores, we computed the difference in games between the participant and their opponent, for every set. For example, a 6–5 loss would be scored as –1; a 6–1 win would be scored as +5. To be able to test our hypotheses, we computed scores that reflected how people performed in decisive versus nondecisive sets.

Most importantly, for each participant, we computed an average game difference (GD) for decisive sets (i.e., 2nd and 3rd sets) and for nondecisive sets (i.e., 1st sets). As choking under pressure is, by definition, a phenomenon that is relative to people's normal performance (Beilock & Gray, 2007), we subtracted these scores ($GD_{23} – GD_1$). We refer to this subtraction as the *choking index* (CI). A higher (positive) CI indicates that people perform better in decisive sets compared with nondecisive sets, whereas a lower (negative) CI indicates that people perform worse in decisive sets. This index, which was computed on the athlete level, was used to provide the main test of our hypotheses.

We further distinguished between different types of decisive sets, to be able to examine the specific performance conditions that athletes encounter during tennis matches. First, we computed a CI specifically for sets in which participants had to win to *not* lose the match, i.e., second sets after the first set is lost ($CI2_{behind} = GD_{2behind} – GD_1$). Second, we computed a CI for sets in which the match could be decided only in the advantage of the participant, i.e., second sets after the first set is won ($CI2_{ahead} = GD_{2ahead} – GD_1$). Third, we computed a CI for sets in which the match can be decided both as a victory and a defeat, i.e., third sets ($CI_3 = GD_3 – GD_1$). These specific CIs are used to attain a more detailed insight into the relationships between BART, AOSPAN, and tennis performance during decisive sets. It should be noted, though, that these specific indices are based on less data points, and are thus less statistically reliable compared with the overall choking index.

Results

Preliminary Analyses: Overall Performance

Although we did not formulate hypotheses about this, we explored whether AOSPAN and BART predicted participants' overall performance. That is, we regressed people's mean game difference (across all sets in the sample, regardless of whether they were decisive) on AOSPAN and BART. In the first step of this regression, only main effects were included in the model. This analyses yielded no main effect of AOSPAN, $\beta = .11$, $t = .6$, $p = .512$, nor of BART, $\beta = .17$, $t = 1.0$, $p = .326$. In the second step of the regression, the AOSPAN × BART interaction was added, which was not significant, $\beta = -.07$, $t = -.4$, $p = .725$. This analysis thus indicated that AOSPAN and BART did not predict participants' overall performance relative to their opponents.

Main Analyses: Choking Index

We tested our main hypotheses by regressing CI on AOSPAN and BART. In the first step of this analysis, only main effects were included in the model. This analysis yielded an effect of AOSPAN, $\beta = .34$, $t = 2.1$, $p = .040$, that indicated that individuals with greater WM capacity performed relatively well in decisive sets. Furthermore, this analysis yielded an effect of BART, $\beta = -.33$, $t = -2.1$, $p = .046$, that indicated that people with higher BART scores (suggesting a more sensitive dopamine system) performed relatively bad in decisive sets. In the second step of the regression, the AOSPAN × BART interaction was added to the model. This interaction proved not significant, $\beta = -.16$, $t = -.90$, $p = .376$. These findings are visualized in Figure 1.

Supplementary Analyses: Choking Index

While all participants were advanced, competitive tennis players, there were still some differences in their level of playing. Moreover, the sample included both men and women. To examine whether the effects of BART and AOSPAN would remain after taking these sources of variance into account, we repeated the above analysis while controlling for playing level (ITN) and gender. This analysis yielded neither significant effects of ITN, $\beta = .11$, $t = 1.2$, $p = .242$, nor gender, $\beta = .51$, $t = 1.5$, $p = .142$. The effect of AOSPAN was still present, $\beta = .36$, $t = 2.3$, $p = .027$. The effect of BART was weaker than in the main analysis, $\beta = -.27$, $t = -1.8$, $p = .090$, but still marginally significant. Like in the main analysis, the AOSPAN × BART interaction was not significant when it was added to the model, $\beta = -.09$, $t = -.49$, $p = .63$. AOSPAN and BART scores correlated neither with ITN, rs $< .02$, ps $> .90$, nor with gender, rs $< .24$, ps $> .17$.

A further issue arose when analyzing accuracy scores of the AOSPAN. Specifically, although participants were instructed to keep their math accuracy higher than 85%, and most participants complied with this instruction (mean accuracy = 91%), six participants did not meet this criterion. Therefore, we repeated the above analyses without these six participants—who could be argued

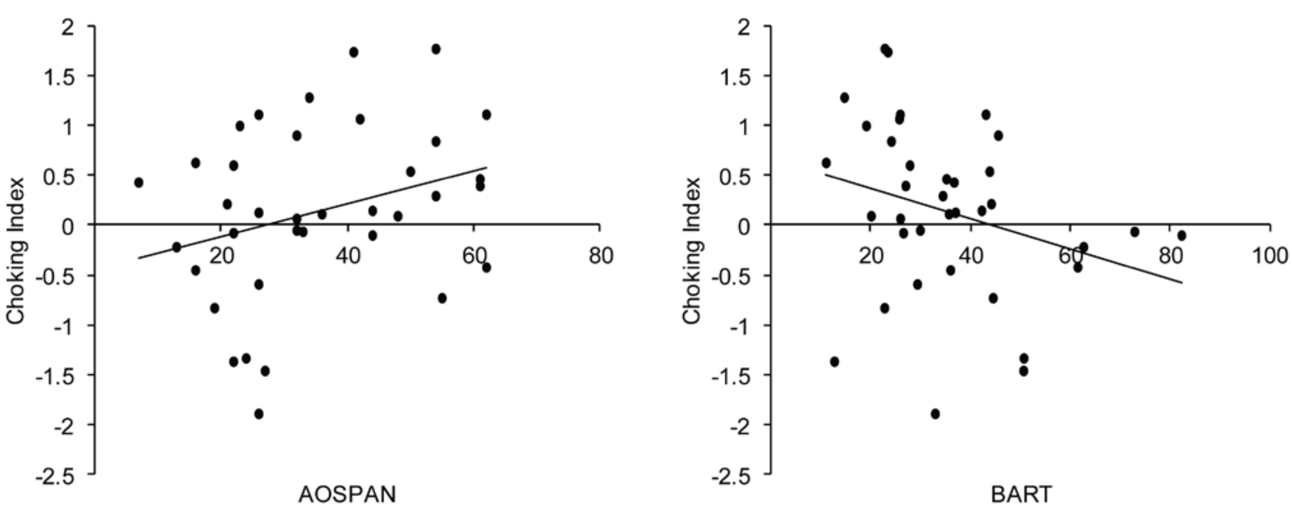

Figure 1 — Scatterplots visualizing the relation between the task scores and tennis performance during decisive sets, as indexed by the choking index. Athletes with a lower choking index tend to perform worse during decisive sets (relative to nondecisive sets).

to have taken the task instructions less seriously—to examine whether their inclusion affected the results. This analysis revealed the same pattern as before: There was a main effect of AOSPAN, $\beta = .41$, $t = 2.2$, $p = .035$, as well as a main effect of BART, $\beta = -.43$, $t = -.2.2$, $p = .033$, but no interaction, $\beta = -.21$, $t = -1.1$, $p = .290$. In this smaller sample, both main effects were still significant after controlling for ITN and gender, AOSPAN, $\beta = .40$, $t = 2.4$, $p = .027$, and BART, $\beta = -.40$, $t = -2.2$, $p = .036$.

Supplementary Analyses: Specific Choking Indices ($CI_{2behind}$, CI_{2ahead}, CI_3)

To explore the effects of AOSPAN and BART on more specific decisive sets, the same regression analyses were done for the specific CIs. For $CI_{2behind}$, main effects of neither AOSPAN, $\beta = .10$, $t = .6$, $p = .575$, nor BART, $\beta = -.14$, $t = -.8$, $p = .412$, were found. However, the AOSPAN × BART interaction was significant, $\beta = -.46$, $t = -2.5$, $p = .016$. Following procedures recommended by Aiken and West (1991), we tested the relationship between AOSPAN and $CI_{2behind}$ separately for people with a high and a low BART score (–1 vs. +1 *SD*). This analysis showed that participants with a low BART score performed better (during decisive sets in which they were behind) when they had a higher AOSPAN score, $\beta = .59$, $t = 2.4$, $p = .025$. However, this relationship was not present for participants with a high BART score, $\beta = -.32$, $t = -1.4$, $p = .17$. In other words, and in line with Hypothesis 3, only participants with low BART scores (suggesting a relatively insensitive dopamine system) performed better due to their greater WM capacity—specifically in sets in which they had to win to remain in the game.

The same analysis was conducted for CI_{2ahead}. This analysis revealed a main effect of BART, $\beta = -.35$, $t = -2.1$, $p = .043$, indicating that people with lower BART scores (suggesting a relatively insensitive dopamine system) underperformed during sets in which they were ahead. The effect of AOSPAN was not significant, $\beta = .06$, $t = .4$, $p = .718$. When the interaction was added to the model, it was not significant, $\beta = -.06$, $t = -.3$, $p = .760$. Finally, the same analysis was conducted for CI_3. The effect of the BART was not significant, $\beta = -.27$, $t = -1.6$, $p = .124$, nor was the effect of the AOSPAN, $\beta = -.02$, $t = -.1$, $p = .930$, nor was the AOSPAN × BART interaction, $\beta = -.11$, $t = -.6$, $p = .557$.

Discussion

The present data support two main conclusions. First, higher-WM (vs. lower-WM) tennis players performed best during decisive sets. This finding supports the idea that people who have greater WM capacity to begin with have less difficulty dealing with the load that is imposed by distracting thoughts that emerge under pressure (Barrett et al., 2004; Vestberg et al., 2012). Second, the current study suggests that tennis players who have a more sensitive dopamine system performed worse during decisive sets. This finding is consistent with research showing that dopamine pathways, when strongly activated, impede PFC functioning and thus thwart performance on various tasks (Arnsten, 2009; Cools & Robbins, 2004). Together, these findings indicate that it is possible to relate real-life performance under pressure to athletes' scores on behavioral tasks that tap into WM functioning. Although performance under pressure is known to be fragile, the present data suggest that athletes with (a) larger WM capacity and (b) a less sensitive dopamine system continue to perform well even during decisive moments.

Results further suggest that for people who have a highly sensitive dopamine system, WM capacity has a smaller impact on performance in decisive sets. Importantly, this finding should be interpreted with caution, as the effect was found only when zoomed in on sets in which the match could be decided only in the disadvantage of the player. Nevertheless, this finding resonates with the idea that the dopamine system can reduce prefrontal control over the rest of the brain, and via that route can reduce the influence of WM on performance (Arnsten, 2009). We can only speculate about why this effect was only visible in this specific type of set. A likely, though post hoc, account for this finding might lie in the idea that sets in which loss is looming are especially pressure-inducing. The match situation may have engaged the dopamine pathways especially strongly during these sets, causing more reliable reductions in the effect of WM.

In general, the supplementary analyses that we conducted to zoom in on specific types of decisive sets, yielded mixed results that are somewhat difficult to interpret. The main effect of BART was in all types of sets negatively related to performance (βs ranging from –.14 to –.35). Although the effect was only significant in one out of three analyses, the consistent pattern leaves room for the interpretation that all types of decisive sets contributed to the general result that BART is negatively related to performance in decisive sets. For AOSPAN, however, the main effect proved less strong in the specific analyses (βs ranging from –.01 to .10, none significant). It is thus possible that the overall main effect of AOSPAN was primarily driven by low-BART athletes' performance during decisive sets in which they could lose the match (i.e., the type of athletes during the type of sets in which AOSPAN was especially predictive of performance, $\beta = .59$). Nevertheless, also here, we stress that findings from these supplementary analyses should be interpreted with caution, as they are based on relatively noisy indicators and as they led to somewhat mixed results. Still, they suggest that it may be interesting and useful for future work to study how the nature of pressure changes when athletes win (or lose) sets, and get closer to (or farther away from) match victory.

While previous research has mostly investigated the psychological underpinnings of choking, the current study suggests that it may well pay off to consider choking under pressure from a neurobiological angle (see Mobbs et al., 2009)—specifically, by looking at the role of the neurotransmitter dopamine. Although we examined the possible role of dopamine indirectly via a behavioral task, which is a limitation of the current study, our findings are in line with previous research showing that dopamine release can thwart performance (Mattay et al., 2003) in the same way as do, for example, very large monetary rewards (Ariely, Gneezy, Loewenstein, & Mazar, 2009). Although the use of dopamine antagonists (i.e., drugs that inhibit dopamine release) is illegitimate in most sports, it would be theoretically interesting to test whether such drugs would enhance performance under pressure specifically for people high in dopamine responsiveness (i.e., people with high BART scores).

Challenging our dopamine-based interpretation, one could argue that high-BART players tend to perform worse during decisive sets not because of dopamine release, but simply because they have the tendency to adopt risk-taking strategies in their tennis game (e.g., hitting second serves really hard, trying risky drop shots and lob shots). We should mention, though, that this alternative explanation requires the assumption that risk-taking behavior is *less* effective during decisive sets than during nondecisive sets. In our view, this assumption is difficult to back up with previous data—that is, some previous research suggests that outcomes in decisive match situations should benefit, not suffer, from risk-seeking choices (Barnett, Reid, O'Shaughnessy, & McMurtrie, 2012). Nevertheless, it would be premature to altogether dismiss a risk-taking-based explanation for the finding that high-BART players perform worse during decisive sets. Specifically, one interesting possibility would be that tennis players who have a strong tendency to take risks (as captured by BART) often try to perform shots that can, when well executed, directly end the rally in their advantage. As such risky shots—e.g., fast second services, drop shots—may often be technically difficult to execute (compared with safer shots), it may be the case that the quality of especially such risky shots suffers under pressure, when people indeed tend to exert dysfunctional amounts of conscious control over the execution of movements (Beilock & Gray, 2007; Masters, 1992). Future research should examine this possibility.

The present study suggests a new way of interpreting tennis scores, by making use of natural variation in decisiveness between sets. The most important advantage of this approach is that it has maximal ecological validity, as performance is recorded in its natural context. Nevertheless, it should be noted that the performance measures are rather coarse, for example, when compared with field research that analyzed specific tennis performance indicators (e.g., service indicators vs. rally indicators; Bijleveld, Custers, & Aarts, 2011). Despite this drawback, the present data indicate that performance during decisive sets can be predicted in ways that are consistent with theories on choking. The present approach to analyzing tennis scores may thus prove fruitful, especially because set scores are much more widely available than full performance statistics (e.g., also from subelite tennis leagues).

Not yet used in the current study, a potential asset of this novel approach is that it can be used to test predictions not only about individual differences, but also about how these individual differences interact with match characteristics (e.g., whether the match is a tournament final, Wells & Skowronski, 2012) and player states (e.g., whether the player is the favorite in a specific match, Jordet, 2009; or whether the player experienced pressure during that specific match, Gucciardi et al., 2010). While this would require more complex statistical models (i.e., multilevel models, Pinheiro & Bates, 2000) and the collection of richer data, the present approach of analyzing tennis scores is potentially useful for uncovering the dynamics of performance under pressure by zooming in to a more detailed level of analysis (e.g., the match level instead of the athlete level). Moreover, we acknowledge that our sample size is relatively small. When taking the present approach, it seems generally desirable to attain a larger sample, especially to convincingly examine the presence and the nature of interactions between multiple predictors (e.g., AOSPAN × BART).

Another limitation of the present research is that it rests on the assumption that pressure is higher in decisive sets, compared with nondecisive sets. As, by definition, the quality of performance is directly connected to the match outcome in decisive sets (and not in first sets), it is likely that athletes experienced these sets to be more important and more stressful. This reasoning resonates with research on test anxiety (Covington, 1985), that suggests that—at least when failure is an option—more important tasks are more likely to be interpreted as a threat, giving rise to negative emotions and cognitions (e.g., worries; Nie, Lau, & Liau, 2011). As we did not collect data on athletes' experiences, however, we cannot exclude the possibility that some athletes experience high pressure also (or perhaps even especially) during first sets. In our view, it would be interesting and important for future research to complement the present approach (i.e., combining task scores and performance data) with self-report measures of how athletes experience match situations (e.g., specific sets in tennis, specific holes in golf).

The present research resonates with the broader idea—widespread in sports science—that aspects of real-life performance can be predicted with tasks and test batteries. While previous psychological research has mainly focused on predicting performance from tasks that measure physical, visuomotor, and general cognitive skills (Blascovich, Seery, Mugridge, Norris, & Weisbuch, 2004; Rampinini et al., 2007; Rowe & McKenna, 2001; Vestberg et al., 2012; Williams, 2000), the current study extends this work toward the prediction of performance under pressure specifically. In so doing, it points to the AOSPAN and the BART as potential candidates for being used in test batteries designed for selection and skill development in sports. Although more research is needed, a future implication may be that athletes low in AOSPAN could benefit especially from some types of psychological skills training. For example, they would likely benefit from interventions known to reduce the negative impact of distracting thoughts (e.g., the mindfulness-acceptance-commitment approach; Gardner & Moore, 2004). People high on the BART, on the other hand, would likely benefit from interventions that help them to get used to and reduce physiological stress (e.g., training sessions under mild pressure; Oudejans & Pijpers, 2010). In so doing, the present research provides a first step into the development of a test battery of tasks that is potentially useful to practitioners in sports settings, helping to design and select interventions tailored to the needs of individual athletes.

Notes

1. A somewhat more detailed characterization of the cascading stress response that leads to activation of the dopamine pathways is as follows: At first, the amygdala engages the hypothalamus and the brain stem, which, in turn, triggers the release of dopamine (Arnsten, 2009). As described in the main text, high levels of dopamine reduce the extent to which the PFC exerts control over activity in other brain structures, including the amygdala (e.g., Goldstein, Rasmusson, Bunney, & Roth, 1996). When the amygdala is no longer under normal PFC control, this may lead to the release of even more dopamine, which may in turn lead to even less PFC control, which may again lead to the release of more dopamine, etc. The dopamine response to stress may therefore result in a cycle that sustains—and perhaps even strengthens—itself, while impairing performance. See Arnsten (2009) for a review on how stress impairs prefrontal functioning via monoaminergic neurotransmitters, including dopamine; see Joëls and Baram (2009) for a broad perspective on the neurobiological underpinnings of stress.

2. Intriguingly, PFC control may also decrease during states of *flow* (i.e., highly focused performance states that are experienced as effortless). Because performance is often increased, not decreased, during flow, we believe the neural underpinnings of flow are vastly different from the neural underpinnings of performance pressure—even though both may involve subnormal PFC control (Dietrich, 2004).

References

Ahmed, L., & de Fockert, J.W. (2012). Focusing on attention: The effects of working memory capacity and load on selective attention. *PLoS ONE, 7,* e43101. PubMed doi:10.1371/journal.pone.0043101

Aiken, L.S., & West, S.G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.

Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. (2009). Large stakes and big mistakes. *The Review of Economic Studies, 76,* 451–469. doi:10.1111/j.1467-937X.2009.00534.x

Arnsten, A.F.T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews. Neuroscience, 10,* 410–422. PubMed doi:10.1038/nrn2648

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews. Neuroscience, 4,* 829–839. PubMed doi:10.1038/nrn1201

Barnett, T., Reid, M., O'Shaughnessy, D., & McMurtrie, D. (2012). Game theoretic solutions to tennis serving strategies. *ITF Coaching and Sport Science Review, 56,* 15–17.

Barrett, L.F., Tugade, M.M., & Engle, R.W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130,* 553–573. PubMed doi:10.1037/0033-2909.130.4.553

Baumeister, R.F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 46,* 610–620. PubMed doi:10.1037/0022-3514.46.3.610

Beilock, S.L. (2010). *Choke: What the secrets of the brain reveal about getting it right when you have to*. New York: Simon and Schuster.

Beilock, S.L., & Carr, T.H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology. General, 130,* 701–725. PubMed doi:10.1037/0096-3445.130.4.701

Beilock, S.L., & Gray, R. (2007). Why do athletes choke under pressure? In G. Tenenbaum & R.C. Eklund (Eds.), *Handbook of Sport Psychology* (3rd ed., pp. 425–444). Hoboken, NJ: Wiley.

Beilock, S.L., Kulp, C.A., Holt, L.E., & Carr, T.H. (2004). More on the fragility of performance: Choking under pressure in mathematical problem solving. *Journal of Experimental Psychology. General, 133,* 584–600. PubMed doi:10.1037/0096-3445.133.4.584

Bijleveld, E., Custers, R., & Aarts, H. (2011). When favourites fail: Tournament trophies as reward cues in tennis finals. *Journal of Sports Sciences, 29,* 1463–1470. PubMed doi:10.1080/02640414.2011.604679

Blascovich, J., Seery, M.D., Mugridge, C.A., Norris, R.K., & Weisbuch, M. (2004). Predicting athletic performance from cardiovascular indexes of challenge and threat. *Journal of Experimental Social Psychology, 40,* 683–688. doi:10.1016/j.jesp.2003.10.007

Buckholtz, J.W., Treadway, M.T., Cowan, R.L., Woodward, N.D., Li, R., Ansari, M.S., . . . Zald, D.H. (2010). Dopaminergic network differences in human impulsivity. *Science, 329,* 532. PubMed doi:10.1126/science.1185778

Cools, R., & Robbins, T.W. (2004). Chemistry of the adaptive mind. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences, 362,* 2871–2888. PubMed doi:10.1098/rsta.2004.1468

Covington, M.V. (1985). Test anxiety: Causes and effects over time. In H.M. van der Ploeg, R. Schwarzer, & C.D. Spielberger (Eds.), *Advances in test anxiety research* (Vol. 4, pp. 55–68). Lisse, The Netherlands: Swets & Zeitlinger.

DeCaro, M.S., Thomas, R.D., & Beilock, S.L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition, 107*, 284–294. PubMed doi:10.1016/j.cognition.2007.07.001

Dietrich, A. (2004). Neurocognitive mechanisms underlying the experience of flow. *Consciousness and Cognition, 13,* 746–761. PubMed doi:10.1016/j.concog.2004.07.002

Driver-Dunckley, E., Samanta, J., & Stacy, M. (2003). Pathological gambling associated with dopamine agonist therapy in Parkinson's disease. *Neurology, 61,* 422–423. PubMed doi:10.1212/01.WNL.0000076478.45005.EC

Eysenck, M.W., & Calvo, M.G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion, 6,* 409–434. doi:10.1080/02699939208409696

Fecteau, S., Pascual-Leone, A., Zald, D.H., Liguori, P., Théoret, H., Boggio, P.S., & Fregni, F. (2007). Activation of prefrontal cortex by transcranial direct current stimulation reduces appetite for risk during ambiguous decision making. *The Journal of Neuroscience, 27,* 6212–6218. PubMed doi:10.1523/JNEUROSCI.0314-07.2007

Fein, G., & Chang, M. (2008). Smaller feedback ERN amplitudes during the BART are associated with a greater family history density of alcohol problems in treatment-naïve alcoholics. *Drug and Alcohol Dependence, 92,* 141–148. PubMed doi:10.1016/j.drugalcdep.2007.07.017

Forbes, E.E., Brown, S.M., Kimak, M., Ferrell, R.E., Manuck, S.B., & Hariri, A.R. (2009). Genetic variation in components of dopamine neurotransmission impacts ventral striatal reactivity associated with impulsivity. *Molecular Psychiatry, 14,* 60–70. PubMed doi:10.1038/sj.mp.4002086

Gardner, F., & Moore, Z. (2004). A mindfulness-acceptance-commitment-based approach to athletic performance enhancement: Theoretical considerations. *Behavior Therapy, 35,* 707–723. doi:10.1016/S0005-7894(04)80016-9

Goldstein, L.E., Rasmusson, A.M., Bunney, B.S., & Roth, R.H. (1996). Role of the amygdala in the coordination of behavioral, neuroendocrine, and prefrontal cortical monoamine responses to psychological stress in the rat. *The Journal of Neuroscience, 16,* 4787–4798. PubMed

Gucciardi, D.F., Longbottom, J.L., Jackson, B., & Dimmock, J.A. (2010). Experienced golfers' perspectives on choking under pressure. *Journal of Sport & Exercise Psychology, 32,* 61–83. PubMed

Jackson, R.C., Ashford, K.J., & Norsworthy, G. (2006). Attentional focus, dispositional reinvestment, and skilled motor performance under pressure. *Journal of Sport & Exercise Psychology, 28,* 49–68.

Jentsch, J.D., Woods, J.A., Groman, S.M., & Seu, E. (2010). Behavioral characteristics and neural mechanisms mediating performance in a rodent version of the Balloon Analogue Risk Task. *Neuropsychopharmacology, 35,* 1797–1806. PubMed

Joëls, M., & Baram, T.Z. (2009). The neuro-symphony of stress. *Nature Reviews. Neuroscience, 10,* 459–466. PubMed

Jordet, G. (2009). When superstars flop: Public status and choking under pressure in international soccer penalty shootouts. *Journal of Applied Sport Psychology, 21,* 125–130. doi:10.1080/10413200902777263

Knutson, B., Wimmer, G.E., Kuhnen, C.M., & Winkielman, P. (2008). Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. *Neuroreport, 19,* 509–513. PubMed doi:10.1097/WNR.0b013e3282f85c01

Kohno, M., Ghahremani, D.G., Morales, A.M., Robertson, C.L., Ishibashi, K., Morgan, A.T., . . . London, E.D. (in press).Risk-taking behavior: Dopamine D2/D3 receptors, feedback, and frontolimbic activity. *Addictive Behaviors*.

Lejuez, C.W., Aklin, W.A., Jones, H.A., Richards, J.B., Strong, D.R., Kahler, C.W., & Read, J.P. (2003). The Balloon Analogue Risk Task (BART) differentiates smokers and nonsmokers. *Experimental and Clinical Psychopharmacology, 11,* 26–33. PubMed doi:10.1037/1064-1297.11.1.26

Lejuez, C.W., Aklin, W.M., Zvolensky, M.J., & Pedulla, C.M. (2003). Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. *Journal of Adolescence, 26,* 475–479. PubMed doi:10.1016/S0140-1971(03)00036-8

Lejuez, C.W., Read, J.P., Kahler, C.W., Richards, J.B., Ramsey, S.E., Stuart, G.L., . . . Brown, R.A. (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology. Applied, 8,* 75–84. PubMed doi:10.1037/1076-898X.8.2.75

Markman, A.B., Maddox, W.T., & Worthy, D.A. (2006). Choking and excelling under pressure. *Psychological Science, 17,* 944–948. PubMed doi:10.1111/j.1467-9280.2006.01809.x

Martins, S.S., Tavares, H., da Silva Lobo, D.S., Galetti, A.M., & Gentil, V. (2004). Pathological gambling, gender, and risk-taking behaviors. *Addictive Behaviors, 29,* 1231–1235. PubMed doi:10.1016/j.addbeh.2004.03.023

Masters, R.S.W. (1992). Knowledge, knerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *The British Journal of Psychology, 83,* 343–358. doi:10.1111/j.2044-8295.1992.tb02446.x

Masters, R.S.W., Polman, R.C.J., & Hammond, N.V. (1993). "Reinvestment": A dimension of personality implicated in skill breakdown under pressure. *Personality and Individual Differences, 14,* 655–666. doi:10.1016/0191-8869(93)90113-H

Mattay, V.S., Goldberg, T.E., Fera, F., Hariri, A.R., Tessitore, A., & Egan, M.F. Weinberger, D. R. (2003). Catechol O-methyltransferase val158-met genotype and individual variation in the brain response to amphetamine. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 6186–6191.

Mesagno, C., & Marchant, D. (2013). Characteristics of polar opposites: An exploratory investigation of choking-resistant and choking-susceptible athletes. *Journal of Applied Sport Psychology, 25,* 72–91. doi:10.1080/10413200.2012.664605

Miller, E.K., & Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167–202. PubMed doi:10.1146/annurev.neuro.24.1.167

Mobbs, D., Hassabis, D., Seymour, B., Marchant, J.L., Weiskopf, N., Dolan, R.J., & Frith, C.D. (2009). Choking on the money: Reward-based performance decrements are associated with midbrain activity. *Psychological Science, 20,* 955–962. PubMed doi:10.1111/j.1467-9280.2009.02399.x

Nie, Y., Lau, S., & Liau, A.K. (2011). Role of academic self-efficacy in moderating the relation between task importance and test anxiety. *Learning and Individual Differences, 21,* 736–741. doi:10.1016/j.lindif.2011.09.005

Nisbett, R.E., & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259. doi:10.1037/0033-295X.84.3.231

Oudejans, R.R.D., & Pijpers, J.R.R. (2010). Training with mild anxiety may prevent choking under higher levels of anxiety. *Psychology of Sport and Exercise, 11,* 44–50. doi:10.1016/j.psychsport.2009.05.002

Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-PLUS*. New York, NY: Springer.

Rampinini, E., Bishop, D., Marcora, S.M., Ferrari Bravo, D., Sassi, R., & Impellizzeri, F.M. (2007). Validity of simple field tests as indicators of match-related physical performance in top-level professional soccer players. *International Journal of Sports Medicine, 28,* 228–235. PubMed doi:10.1055/s-2006-924340

Rao, H., Korczykowski, M., Pluta, J., Hoang, A., & Detre, J.A. (2008). Neural correlates of voluntary and involuntary risk taking in the human brain: An fMRI study of the Balloon Analogue Risk Task (BART). *NeuroImage, 42,* 902–910. PubMed doi:10.1016/j.neuroimage.2008.05.046

Rowe, R.M., & McKenna, F.P. (2001). Skilled anticipation in real-world tasks: Measurement of attentional demands in the domain of tennis. *Journal of Experimental Psychology. Applied, 7,* 60–67. PubMed doi:10.1037/1076-898X.7.1.60

Sorg, B.A., & Whitney, P. (1992). The effect of trait anxiety and situational stress on working memory capacity. *Journal of Research in Personality, 26,* 235–241. doi:10.1016/0092-6566(92)90041-2

Unsworth, N., Heitz, R., Schrock, J., & Engle, R. (2005). An automated version of the operation span task. *Behavior Research Methods, 37,* 498–505. PubMed doi:10.3758/BF03192720

Vestberg, T., Gustafson, R., Maurex, L., Ingvar, M., & Petrovic, P. (2012). Executive functions predict the success of top-soccer players. *PLoS ONE, 7,* e34731. PubMed doi:10.1371/journal.pone.0034731

Wang, J., Marchant, D., Morris, T., & Gibbs, P. (2004). Self-consciousness and trait anxiety as predictors of choking in sport. *Journal of Science and Medicine in Sport, 7,* 174–185. PubMed doi:10.1016/S1440-2440(04)80007-0

Wells, B.M., & Skowronski, J.J. (2012). Evidence of choking under pressure on the PGA Tour. *Basic and Applied Social Psychology, 34,* 175–182. doi:10.1080/01973533.2012.655629

White, T.L., Lejuez, C.W., & de Wit, H. (2008). Test–retest characteristics of the Balloon Analogue Risk Task (BART). *Experimental and Clinical Psychopharmacology, 16,* 565–570. PubMed doi:10.1037/a0014083

Williams, A.M. (2000). Perceptual skill in soccer: implications for talent identification and development. *Journal of Sports Sciences, 18,* 737–750. PubMed doi:10.1080/02640410050120113

Manuscript submitted: March 11, 2013

Revision accepted: April 15, 2014