Summer 2016

Are there bad wins and good losses? Outcome effects in sport

Richard Erik Inglis
James Madison University

Are There Bad Wins and Good Losses? Outcome Effects in Sport

Richard Erik Inglis

A thesis submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Master of Arts

Psychological Sciences

August 2016

FACULTY COMMITTEE:

Committee Chair: Dr. Robert Harmison

Committee Members/ Readers:

Dr. Kethera Fogler

Dr. Trevor Stokes

Acknowledgments and Dedication

The author would like to thank Dr. Kethera Fogler for her helpful suggestions on how to improve this manuscript. Thank you to my advisor, Dr. Robert Harmison, who helped shape my ideas and my writing in our innumerable meetings. This work is dedicated to my parents, Pia Eriksson-Inglis and Evan Inglis, in appreciation for all their support and direction throughout the years. To my grandfather, Dick Inglis, whose tableside chats helped inspire my inquisitive nature. And finally, to Elisabeth Huber, whose unstinting encouragement and tireless love is a large reason why I was able to complete this project. My hope is that this work is worthy of your praise.

Correspondence concerning this article should be addressed to Richard Erik Inglis, College of Education, Florida State University, Stone Building, 1114 W. Call Street, Tallahassee, FL, 32306. Email (preferred): Inglisre@dukes.jmu.edu *or* Rei16@my.fsu.edu

Table of Contents

List of Tables

Abstract

Outcomes have been shown to distort feeling states, performance evaluations, and subsequent performances. This retrospective distortion represents an outcome bias, a phenomenon that has been studied extensively in a wide array of disciplines and domains. However, despite their importance, the effects of outcomes have received little attention in the sport corpus. In an attempt to remedy this oversight, 60 participants with basketball experience were recruited to take part in a free throw shooting competition. Participants completed two rounds of 20 free throws. Following their first round, participants were randomly assigned to either a negative outcome (lose by one shot) or a positive outcome (win by one shot). Subsequently, participants were asked to reflect on their first round performance and to respond to a modified version of the CSAI-2 and a performance evaluation item. It was hypothesized that the positive outcome group would rate their performance as better, report lower levels of cognitive anxiety and higher levels of self-confidence, and improve their score in a subsequent shooting trial. In the main, a series of *t*-tests revealed that these hypotheses were not supported. However, a Cohen's *d* test revealed a medium effect of cognitive anxiety in the hypothesized direction. Further, a Cohen's *d* test found that assigned outcome had a medium effect on future performance. Males who received a positive outcome following their round 1 performance in turn performed stronger in round 2 and vice versa. In addition, several correlations and means bear mention. These results are suggestive of an outcome bias. In light of these findings, implications for practitioners and other sport professionals are offered and lines of future research are recommended.

Keywords: outcome bias, sport, performance evaluation, feeling states

Are There Bad Wins and Good Losses? Outcome Effects in Sport

Have you ever stood 3 feet away from writing your name in the history books? Take a step into the cleats of quarterback Russell Wilson and his Seahawk teammates as Russell hovered above center, Max Unger, and surveyed the field inches away from the goal line and glory. The Seattle Seahawks are entrenched in a battle with the New England Patriots in football's biggest game, the Super Bowl. One yard and 26 seconds is all that stands between Seattle and back-to-back Super Bowl titles.

What transpired in the succeeding seconds will forever be one of the most discussed plays in football. In lieu of handing the ball off to Marshawn Lynch, who is well known for his truculent running style, Seattle opted to throw. Russell dropped back and slung the ball to wide receiver Ricardo Lockette at the goal line, but Ricardo never received the delivery. Instead, Malcolm Butler, a cornerback for the Patriots, intercepted the ball thereby ending the game and forever cementing himself in football lore. Just like that, all of the air left the ball for the Seahawks. Their ignominious fate that day is now the stuff of sport legend, and in the days following the Super Bowl many chided the final play call. So great was the disdain for this decision that Emmitt Smith, the revered former running back for the Dallas Cowboys, likened it to the worst play call ever (Puckett, 2015). Emmitt was not alone in his sentiment. In fact, it wasn't just outside observers who were left scratching their heads. Even some Seahawk players were nonplussed by the decision. Linebacker Bruce Irvin, still struggling to come to terms with the result, wondered aloud, "We had it. I don't understand how you don't give it to the best back in the league, on not even the one-yard line" (Reyes, 2015).

Despite the excoriating language, statistics show that the final play call was not as straightforward as some believed it to be. Consider that prior to the Seahawks' final play there were 334 examples of plays run from the one-yard line during the 2014-2015 NFL season (Blunt, 2015; Caplan & Igel, 2015; Morris, 2015). In 219 (65.6%) of these examples teams elected to run the ball, while in 115 (34.4%) teams elected to pass the ball. All told, 70 throws (60.9%) from the one-yard line resulted in touchdowns, while 125 runs (57.1%) resulted in touchdowns yielding a slight edge to passes percentagewise. Moreover, 23 of these run plays (10.5%) resulted in a loss of yards. Still more damning is the fact that twice that year run plays from the one-yard line resulted in catastrophic failure (i.e., a fumble recovered by the opponent), while no interceptions were thrown. To be sure, the debate as to whether or not the Seahawks should have run the ball is not without merit, but the notion that it was the worst play call ever should be considered risible. So, what then was the cause for all the vitriol thrown the Seahawks way?

"Winning isn't everything, it's the only thing" (Overman, 1999, p. 77). This quote from the venerated former NFL football coach, Vince Lombardi, highlights the importance placed on winning (i.e., outcome) in sport. When Russell Wilson threw that fabled interception, soiling the Seahawks chances at joining a rare class of back-to-back champions, all that seemed important to the pundits was the fact that the Seahawks lost the game. Indeed, often in sport the outcome-performance relationship is portrayed apodictically. That is, a good result necessarily follows a good performance. However, this is not always the case, and there is good reason to question the putative strength of this relationship. Accordingly, when emphasis on the outcome is too strong, performance evaluators (e.g., athletes, coaches, fans) may improperly associate a poor outcome with a

poor performance (Lefgren, Platt, & Price, 2013; Plessner & Haar, 2006) and vice versa. This phenomenon represents an outcome bias (Hershey & Baron, 1988) or outcome effects (these terms will be used interchangeably throughout this article), where undue weight is set to results when evaluating performance. Outcome bias nearly resembles hindsight bias (Roese & Vohs, 2012), but is demonstrably unique in that remembrances of the past do not necessarily favor the actor. Thus, harkening back to the foregoing anecdote, when individuals (even experts) evaluate the Seahawks' final play call, and Seattle's performance on a whole, evaluators' perceptions are vitiated by the play's hapless outcome.

Inherit to all competition, a certain degree of luck accompanies any given sporting event. This means that an athlete, coach, or team can perform optimally and not be successful (i.e., win) and vice versa. In such a scenario, it is imperative that the optimal performance be the focal point, not the outcome, to increase the likelihood of future success (Lefgren, Platt, & Price, 2013). If outcomes are championed in spite of performance, athletes and coaches may set themselves up for failure in subsequent competition. Therefore, it is incumbent upon performance evaluators to be wary of the potential deleterious effects of an outcome bias. Despite the fact that outcome effects have been studied extensively in fields and domains such as judgment and decision-making (Baron & Hershey, 1988), medicine (Sacchi & Cherrubini, 2004), accounting (Mertins, Salbador, & Long, 2013), ethical reasoning (Gino, Moore, & Bazerman, 2009), military strategy (Lipschitz, 1989), and criminal justice (Mazzocco, Alicke, & Davis, 2004), the effects of outcome have received little attention in the sport literature.

Therefore, the purpose of this study is to determine whether outcomes play an elemental role in performance judgments and future performance in sport and, if so, to what extent.

Statement of purpose. The goals of this review and research are as follows: first, the need in sport for understanding how outcomes affect performance evaluations is highlighted. Next, previous research on outcome effects both in and outside sport will be explored. Specifically, the potential for outcomes to obscure evaluations of performance, distort reflections of perceived cognitive anxiety and self-confidence, and in turn tamper with future performance, will be evaluated. Subsequently, a procedure for testing the effects of outcomes in sport (i.e., basketball) is outlined. Next, a description of the participants and the measures used in this experiment is provided, followed by an analysis of the results. In addition to the main hypotheses (introduced later), a series of correlations (e.g., self-confidence and performance), potential gender differences, and the possibility of skill level effects are also investigated. Finally, a discussion of the results as they relate to the hypotheses, the potential significance of these findings, and limitations are all explored in detail.

The Need for More Objective Measures of Performance

Recently, there has been a push to find more objective measures of performance in sport (Drust, 2010). Coaches, athletes, and sport scientists all consider performance analysis to be integral to their success. Advances in computer and video technology, such as multicamera tracking systems, have afforded performance analysts with the ability to quantify performance variables and provide coaches and athletes with a complex measure of achievement. For instance, a notational analyst (NA) who is evaluating the performance of a soccer team may be interested in quantifying the number of shots,

passes, and passing accuracy. In a similar vein, for tennis, a NA may wish to focus on quantifying winners-to-errors ratios, shots/rallies, and quality serves and returns. Despite statistical analyses providing coaches and athletes with sought-after breadth, qualitative analyses of performance are far from obsolete (Nelson & Groom, 2012). Indeed, coaches and athletes often engage in discussions about performance sans the help of numbers. If qualitative evaluations of performance are still central to the athletic endeavor, analysts must be wary of potential biases in their evaluations. Therefore, one goal of the current study is to further understand how outcomes retrospectively distort perceptions of performance. Such knowledge will allow coaches, athletes, and other interested parties to take a more objective approach to their assessments.

Review of Literature

Outcome Effects Outside of Sport

Outcome bias and public opinion. Research in several domains has explored the effects of outcome on perceptions of decision-making and performance. Specifically, in the study of judgment and decision-making, outcomes have been reviewed for nearly seventy-five years (Kahneman & Tversky, 1986). Fischoff (1975) observed that knowledge of outcomes increased judgments of the perceived likelihood of an event occurring. Subjects were presented with an arcane historical event (e.g., the British-Gurka struggle in India in the early 1800s) and asked to judge the likelihood of four potential results. Individuals were randomly assigned to either a before-outcome or after-outcome group. In the after-outcome group, participants were randomly assigned to one of the four results (e.g., the British won) and were told that this was the actual result of the event. Subsequently, participants were asked to judge the likelihood of all four

possible results occurring—those in the after-outcome group were asked to disregard their assigned result. Nonetheless, participants in the after-outcome group rated their provided result, on average, as 10.8% more likely than the before-outcome group. Furthermore, individuals in the after-outcome group were more likely to rate their result as inevitable (subjects in the before-outcome group did not rate any of the possible results as inevitable) and the other results as impossible. Evidently, the result of an event looms heavy and it is difficult to disregard its saliency. Thus, it is challenging to paint a valid picture of an event as it developed when outcomes are known.

Additionally, Baron and Hershey (1988) witnessed an outcome bias when individuals rated decisions made in a fictional medical matter. Participants were presented with a case where a 55-year old man had a heart condition and a decision had to be made on whether or not to operate. Raters were informed that the operation resulted in death in 8% of surgeries. Researchers manipulated both the decision maker (i.e., a physician or the patient) and the result of the procedure (i.e., success or failure). Graders exhibited an outcome bias by rating the decision to operate with successful outcomes (i.e., the patient lives) as better than those with unsuccessful outcomes (i.e., the patient dies) despite the fact that the odds of failure were the same for both groups (8%). This result illustrates that the judgers misconstrued the outcome as informative to judgments of the decision-making process. Moreover, when the outcomes were positive, the decision-makers were rated as better thinkers and more competent by a third-party reviewer.

Similarly, outcome effects were seen in a study that assessed the perceived culpability of an actor (i.e., homeowner) in a fabricated aggressive act (Mazzocco et al.,

2004). Participants were asked to judge how blameworthy and negligent the homeowner was in an attack on a perceived intruder (i.e., armed robber, homeowner's daughter's boyfriend, or no indication) as well as how long the homeowner should be sentenced to prison. When the intruder was only injured, the blame results were significantly varied depending on the weapon of choice (i.e., gun or bat). However, when the intruder was killed as a result of the attack, no significant differences were observed based on the weapon of choice.

The afore examples all deal with hypotheticals, but Fischoff and Beyth (1975) demonstrated the real world applicability of an outcome bias when they asked a group of university students to predict the probabilities of certain outcomes for a planned event. Participants judged the likelihood of certain outcomes (e.g., President Nixon will meet Mao at least once) for President Nixon's upcoming trip to Peking and Moscow. The unwitting students were then asked to recall their original probabilities a few weeks after the trip was complete. The group exhibited an outcome bias by offering greater probabilities for the events they felt had occurred and lower probabilities for the ones they believed had not occurred. The preceding research illustrates that outcomes can bias performance evaluations; however, in none of these studies were the participants intimately familiar with the subject matter at hand.

Experts and performance evaluations. Lipschitz (1989) surveyed the opinions of a group of cadets at an Israeli military school on a contrived military matter. Cadets were provided with a passage that detailed a wartime scenario wherein a soldier had to decide whether to follow protocol (i.e., appropriate action) or disobey (i.e., inappropriate action). The passage concluded by providing an outcome (successful or unsuccessful)

following the soldier's decision to go with or against protocol. Hence, the cadets were assigned to one of four potential groups (appropriate-successful, appropriate-unsuccessful, inappropriate-successful, inappropriate-unsuccessful). On the whole, following a successful mission, cadets rated the hypothetical soldiers decision as the product of more scrupulous evaluation of the consequences as well as more appropriate. Further, soldiers who found success were rated as more active than passive, more decisive, more deliberate, more justified in their response, more responsible for the result, and as having more initiative. Thus, even those dealing with a topic they are learned in can be biased as a consequence of outcomes.

Disconcertingly, outcome effects likely play a role in performance evaluations in the work place (Marshall & Mowen, 1993). A group of marketing students rated the performance of a hypothetical salesman as stronger following a successful sale (i.e., a positive outcome). Students exhibited an outcome bias by offering strong evaluations even when the method used by the salesman was considered dodgy and the eventual sale fortuitous. This is unsettling because it suggests that performance evaluations may simply be a reflection of sales performance, rather than a review of genuine performance. Taken together, these results further suggest that outcomes have an effect on the evaluations of others' performance, but what about evaluations of one's own performance?

Personal performance evaluations. Personal perceptions of judgments have shown to be retrospectively distorted by the knowledge of outcomes. For instance, Jones, Yurak, and Finsch (1997) found that perceptions of judgment were biased when a group of participants were presented with a hypothetical scenario in which they had to choose between two comparable scholarship applicants. Participants were provided with a

manipulated outcome (i.e., correct selection, wrong selection, or no outcome) and then asked to grade their decision-making process. When asked to rate how confident they were in their choice, participants who chose correctly rated their performance significantly better than both those in the no-outcome group and individuals who selected the wrong applicant. Furthermore, when they were asked to rate the quality of their decision-making process, those who chose correctly again rated their decision-making more favorably than the negative outcome group (there was no significant difference between the no outcome group and the positive outcome group).

Likewise, Sacchi and Cherrubini (2004) examined outcome effects among doctors. A group of Italian physicians were presented with a mock clinical case that entailed 8 distinguishable symptoms. Unbeknownst to the physicians, 4 of the symptoms were related to Lewy body syndrome and 4 were related to alcohol withdrawal. The physicians were asked to provide a diagnosis (either Lewy body syndrome or alcohol withdrawal) and, despite there being no correct diagnosis, were then randomly assigned to either a good-outcome (i.e., correct diagnosis), no outcome, or bad-outcome group (i.e., incorrect diagnosis). Those in the positive-outcome group rated their decision-making process as significantly better than those in the other two groups. Furthermore, the positive-outcome group rated their decision-making process as far less difficult than the bad-outcome group (there was no significant difference between the no outcome group and positive-outcome group). These results elucidate that knowledge of a consequence can warp our opinions of personal decisions.

Outcomes and future performance. Outcomes also have the power to affect future performance (Denes-Raj & Epstein, 1994; Erev & Roth, 1998; Ratner & Herbst,

2005). Erev and Roth (1998) noted results analogous to the aforementioned studies in an analysis of gaming strategy. They found that positive outcomes had significant predictive power of future performance. That is, when the results were positive, individuals were likely to implement that same strategy in subsequent trials of the gaming task. This implies that positive outcomes are seen as representative of optimal performance, and therefore, reinforce behavior.

In addition, Denes-Raj and Epstein (1994) found that poor outcomes lead to switching strategies. Participants in a decision-making task were presented with the chance to win a prize if they selected the winning color from two trays of jellybeans (Denes-Raj & Epstein, 1994). One tray had a greater proportion of winning jellybeans (10% compared to 7%) and therefore should have been selected on each trial. However, despite the fact that one tray yielded a greater chance of success, 80% of participants selected a jellybean from both trays. When the optimal tray did not yield a winning jellybean, participants were swayed to select from the other, less plentiful tray (i.e., performance was affected).

Ratner and Herbst (2005) found that these results generalized to a hypothetical investment task. In this study, participants were asked to select between broker A, who had a 54% chance of making a successful investment, and broker B, who had a 43% chance of making a successful investment (Ratner & Herbst, 2005). Initially, all participants selected broker A. If participants learned that broker A was successful, 98% of them stayed with broker A in a subsequent trial. However, if they learned that the broker was not successful, only 77% of participants stuck with broker A. Interestingly, even when initial probabilities of success are made salient, this switching behavior is not

mitigated, revealing the robustness of this phenomenon (Arkes, Dawes, & Christensen, 1986; Ratner & Herbst, 2004). These results signal that an outcome can lead us to perform in a suboptimal fashion in future trials of an event.

Outcome Effects in Sport

A dearth of research exists exploring outcomes (e.g., wining and losing) and their affect on performance evaluations and future performance in sport. Despite the scant studies that exist, a few provide relevant implications. For instance, Plessner and Haar (2006) found that the knowledge of results could affect evaluations of sport performance. When a group of novice and expert judgers were asked to grade the performance of a soccer team after viewing a clip of said team's game, the judgers' perceptions were skewed by a manipulated outcome (i.e., the team won or the team lost). Both the novice and expert evaluators who were told that the team won rated their performance better than the evaluators who were told that the team lost. Ostensibly, it makes sense that outcomes would heavily skew novice's judgments because they lack experience with the sport. However, even expert judgers were not immune to an outcome bias. Indeed, Plessner and Haar (2006) found that expert evaluators were more susceptible to a bias than non-experts. That is, there was a greater discrepancy between the performance ratings of the experts in the winning group and the losing group than the discrepancy among novice performance ratings.

Arkes et al. (1986) observed corresponding results in a group of moderately and highly knowledgeable baseball fans. When participants were asked to predict who won the Most Valuable Player (MVP) award from 1940-1961 based on 7 statistics (e.g., number of home runs hit, runs batted in, etc.), the highly knowledgeable group had

significantly fewer correct responses despite being more confident in their responses. Thus, experts are arguably at a greater risk for an outcome bias.

Outcome bias has been recognized on the playing field as well. Three economists (Lefgren et al., 2011) explored the effects of outcomes on coaches and their strategy (i.e., the plays they run) in 5,661 NFL games over 25 seasons (1985-2009). Their results demonstrated that coaches become complacent (i.e., do not change strategy) following wins and alter strategy excessively following losses. Strategy was measured by a team's pass-to-run ratio, and on average NFL teams threw the ball 53.7% (*SD* = 11.6%) of the time over this period. On the whole, passing plays lead to 2.6 more yards than running plays, but a pass is still seen as riskier than a run because the quarterback may throw an incomplete pass and gain no yards, get sacked and lose yards or, worse, throw an interception. Consequently, pass frequency was abated following losses, but maintained following wins. Said another way, coaches were more likely to throw the ball more in games succeeding wins than they were in games succeeding losses. Curiously, Lefgren et al. (2011) found that the magnitude of a loss had no impact on the extent to which strategy was manipulated. That is, a coach that lost by 20 altered their strategy to the same extent as a coach that lost by two.

Analogous results were seen in a 19-year (1991-2010), 46,550 game review of NBA coaching decisions. Following losses, coaches were more likely to adjust their strategy. Ostensibly, these coaches felt that a loss was a rebuke of their initial stratagem. In response, they tended to alter their starting lineup in a subsequent game. Coaches who won, however, were inclined to stay with their starting lineup. Both in the NBA and in the NFL, strategy manipulations were not dependent upon the magnitude of the win or

the loss despite the fact that close results have no predictive utility for future performance. Further, expected performance did not factor into the equation. For instance, even when a team was projected to lose yet they beat the point spread (i.e., expected magnitude of loss) and outperformed expectations, coaches still responded with change.

Remarkably, Berger and Pope (2011) found that professional and collegiate basketball teams that were slightly behind at halftime were more likely to win the game than would be expected. In fact, teams that were down by one at halftime were more likely to win the game. Thus, the argument could be made that losing is sometimes valuable and a boon for performance. Interestingly, Arkes et al. (1986) found that financial incentives exacerbated outcome effects in a decision-making task. It may be that professional (i.e., paid) coaches and athletes are at a greater risk for outcome effects and Lefgren et al. (2011, 2014) make explicit note of this distinct possibility. Furthermore, for management, fans, and the media, which often play an integral role in the employment of professional coaches, outcomes may render performance evaluations obtuse. This is a troubling notion for coaches whose jobs are, arguably unfairly, directly dependent upon results (Lefgren & Platt, 2010).

Self-confidence and Cognitive Anxiety in Sport

Ratner and Herbst (2005) illustrated how negative emotional responses to outcomes can lead individuals to stray from optimal decision-making in a hypothetical investment example. In sport, two affects that are related to success and that may be skewed as a result of outcomes are self-confidence and cognitive anxiety (Woodman &

Hardy, 2003). Speaking to the importance of confidence, a pentathlete offered this insight regarding self-confidence and fencing:

> When you're confident on the piste you get your distances much better . . . When you're not confident in what you're doing you tend to be more timid in your movement and the thing with fencing is, once you go for a move you've just got to go for it (Hays, Thomas, Maynard, & Bawden, 2009, p. 99).

A meta-analysis of 48 studies revealed that both self-confidence (r = .24) and cognitive anxiety (r = -.10) were related to performance (Woodman & Hardy, 2003). Furthermore, the relationship was greater for higher-standard athletes (e.g., elite athletes) than lower-standard athletes. One potential reason for this phenomenon is that high-standard athletes are under greater pressure and stress than their low-standard peers. Further, Jones, Hanton, and Swain (1994) found that elite athletes (specifically swimmers) exhibited more self-confidence than non-elite athletes.

Despite the results of the aforesaid meta-analysis, there is no consensus yet among sport researchers, as these relationships (i.e., performance and self-confidence, cognitive anxiety) have tallied directional discrepancies. For instance, mirroring the results of the meta-analysis, Martin and Gill (1991) reported that self-confidence (r = .57) was a significant predictor of performance in a group of high school middle and long distance runners, but that cognitive anxiety (r = -.10) was not. Conversely, Edwards and Hardy (1996) found the inverse in a group of elite female netball players where self-confidence negatively predicted performance while cognitive anxiety was positively

related to performance. Likewise, Taylor (1987) found that cognitive anxiety positively predicted performance in university athletes in five different sports. Contrarily, Jones, Swain, and Hardy (1993) found that self-confidence ratings predicted future success in a group of teenage gymnasts, but that cognitive anxiety did not. These are not the only discrepant findings in the corpus of sport research. Indeed, Woodman and Hardy (2003) promulgated dozens of studies that offered disparate findings in a wide range of sports (e.g., swimming, pistol shooting, basketball, and artistic gymnastics).

In addition, gender may play a significant role in these relationships. Consider that in Woodman and Hardy's (2003) meta-analysis, both self-confidence and cognitive anxiety significantly predicted performance in men while neither relationship was significant for women. Further, how and when feedback is provided may also be a principal factor for performance. Chiviacowsky and Wulf (2007, 2009) provided participants in a beanbag-tossing exercise with feedback following either their most accurate or least accurate throws. Subsequently, both groups repeated the task, and those who received feedback following their most accurate throws performed stronger. Likewise, Hutchinson, Sherman, Martinovic, and Tenenbaum (2008) found feedback relative to competitors to be a central predictor of future performance. Participants were asked to squeeze a dynamometer with 25% of their strength for as long as they possibly could. When participants reached exhaustion they were provided with arbitrary results that indicated that they performed either in the bottom 10% (i.e., low self-efficacy group) or top 10% (i.e., high self-efficacy group) of all participants. These farcical results were shown to impact performance on subsequent trials such that those in the low self-efficacy group performed worse than they previously had and vice versa. Furthermore, when

participants were invited to judge their performance relative to others, this phenomenon was even more impressive. These results align with the rationale put forth by Hanton, Mellalieu, and Hall (2004), who posited that athletes with lower self-confidence tend to recall poorer performances while the inverse is true of athletes with greater self-confidence. If outcomes can skew perceptions of performance (e.g., Hutchinson et al., 2008; Sacchi & Cherubini, 2004) they may have similar affects on self-perceptions of cognitive anxiety and self-confidence. Specifically, negative outcomes may attenuate perceptions of self-confidence and augment perceived cognitive anxiety.

Literature Review Summary

To recapitulate, in sport, a premium is placed on outcomes. Namely that winning is indicative of optimal strategy while losing is indicative of poor strategy. In response to poor outcomes, as witnessed in the Super Bowl anecdote, pundits have a proclivity for prosaic narratives replete with undue scorn and resentment over the result. This is particularly troublesome when you consider that results are not always in line with optimal strategy (Lefgren et al., 2011, 2014). Accordingly, there has been a call in sport research for more objective measures of performance (Drust, 2010).

Though only a modicum of research exists exploring the effects of outcomes in sport, outcomes have a more studied history in other domains including judgment and decision-making (Baron & Hershey, 1988) and medicine (Sacchi & Cherubini, 2004). This literature suggests that being knowledgeable on a subject does not negate or attenuate any potential bias (Lipschitz, 1989; Marshall & Mowen, 1993). Further, when judging the performance of others, as seen in the hypothetical 55 year-old heart patient and the conjured operation scenario (Baron & Hershey, 1988), outcomes can skew

evaluations of performance. Likewise, when assessing our own decisions, outcomes can slant our evaluations as suggested by the Italian physicians and the diagnosis decision (Sacchi & Cherrubini, 2004). In turn, future performance may also be altered on the basis of outcomes. As mentioned, Ratner and Herbst (1994) found that participants made suboptimal investment choices in a hypothetical investment scenario when an optimal choice failed to yield desired results.

More locally, an outcome bias has been found in the sport corpus. Even among experts, performance evaluations of a soccer team's highlight film were retrospectively distorted with knowledge of the result (Plessner & Haar, 2006). That is, when evaluators were informed that the team being evaluated won, they offered more favorable appraisals of their skill than if they were told this team lost. Similarly, in a 25-year review of NFL play calling, Lefgren et al. (2011) found that coaches tended to, magnitude of win or loss notwithstanding, manipulate their schemes following losses. On the other hand, they tended to stay the course strategically following wins.

Furthermore, both self-confidence and cognitive anxiety, as was the case in the dynamometer exercise (Hutchinson et al., 2008), may be influenced by results. In a meta-analysis, Woodman and Hardy (2003) found that the impact of outcomes on perceptions of cognitive anxiety and self-confidence was significant, but only for males. In an attempt to deepen the understanding of the effects of outcomes in sport, the successive experiment, results, and discussion are offered.

The main objective of this research is to determine if outcomes will pervert performance evaluations and warp perceptions of cognitive anxiety and self-confidence. Furthermore, this research will investigate the impact of outcomes on future performance,

explore potential gender differences, and evaluate the effects of skill level. An ancillary aim of this research and discussion is to ignite serious conversation amongst actors in sport about the pernicious effects of embracing a myopic focus on outcomes. Lastly, the current study is unique from previous research in that the effects of outcomes are tested empirically in a sport setting. After manipulating outcomes, performance is measured to determine if results alter future performance.

Method

Participants

Participants were 49 male students and 11 female students from James Madison University who self-selected participation through an online database (SONA) in order to meet a requirement set forth by the institution. In addition, a few participants were recruited through a flyer emailed to all athletes in James Madison's club sports. As a prerequisite, all athletes must have competed on an organized basketball team while in high school (9^{th}–12^{th} grade). Participants represented a wide variety of skill levels. In total, 26 indicated recreational basketball as being their highest level of play, five indicated freshman basketball as their highest level of play, 13 indicated junior varsity, 14 indicated varsity, and one indicated club basketball at the university level as being their highest level of play. The majority of participants (M_{age} = 19.5 years, SD = 1.03 years, age range: 18-23 years) were freshmen (n = 32), followed by sophomores (n = 16), seniors (n = 6), juniors (n = 5), and graduate students (n = 1). Further, the bulk of participants were Caucasian (n = 49), followed by Asian (n = 6), then African-American (n = 3), next Other (n = 2), and finally Native Hawaiian or Pacific Islander (n = 1). All

were assured anonymity, and informed consent was obtained prior to participation. Participants were compensated $5 for their involvement.

Measures

CSAI-2. A modified version of the Competitive State Anxiety Inventory–2 (CSAI-2; Martens, Vealey, & Burton, 1990) was administered during the current study. The CSAI-2 is a 27-item questionnaire that measures self-perceived competitive state anxiety on three subscales: cognitive anxiety, somatic anxiety, and self-confidence. An example of an item testing cognitive anxiety is "I am concerned about losing." An example of an item testing somatic anxiety is "I feel jittery." An example of an item testing is self-confidence is "I feel secure." Each subscale contains nine items on a four-point Likert scale with total scores ranging from 9–36. This measure is considered highly reliable as indicated by Cronbach's alpha with reliability coefficients ranging from α = .79, .83, α = .82, .83, and α = .87, .90, for the cognitive anxiety, somatic anxiety, and self-confidence subscale, respectively (Ostrow, 2002). The CSAI-2 has been used to examine competitive state anxiety in numerous sports including basketball, cycling, golf, swimming, track & field, and wrestling (Woodman & Hardy, 2003). Indeed, use of the CSAI-2 is still ubiquitous in current sport literature. Recently, the CSAI-2 was administered on a group of elite basketball players in order to test the influence of home and away games on state anxiety (Arruda et al., 2014). At the current, only cognitive appraisals associated with performance were of interest. Thus, for the purposes of the current study only the cognitive anxiety and self-confidence subscales were utilized.

Self-perceived performance. For the purposes of the current study an item was created to measure the participants' self-perceptions of their performance. Specifically,

participants were asked to recall and grade their performance by responding to the following question: “How would you rate your performance?” This item requires participants to provide a self-perceived rating of performance from 1 (very bad) to 7 (very good). This inquiry was adapted from a question used in Sacchi and Cherubini’s (2004) study on outcome effects in doctors. In their version, the question reads: “How well would you rate your performance?” For the current study, “well” was dropped because there were concerns that this was a leading question (Whitney & Kite, 2012).

Post - performance improvement. For the purposes of the current study an item was created to measure the participants’ self-perceptions of the need for improvement in their performance. Specifically, participants were asked to recall their performance and indicate to what degree they felt they needed to improve in a subsequent performance by responding to the following query: “Compared to your first performance, how do you feel you must perform in the second performance?” This item required participants to provide a self-perceived rating of need for improvement ranging from 1 (much worse) to 7 (much better). Ultimately, this item was dropped from analyses as several participants indicated confusion over the questions wording.

Exit survey. At the conclusion of the experiment, participants were asked to complete a brief exit survey. This survey was intended to determine whether the participants in both groups provided commensurate effort and derived similar enjoyment, as well as to allow participants an opportunity to provide feedback. The first item queried, “To what extent did you find this experiment enjoyable?” This item required participants to indicate their enjoyment on a scale of 1 (not at all enjoyable) to 5 (very enjoyable). The second item queried, “To what extent did you take this task seriously?”

This item required participants to indicate their effort on a scale ranging from 1 (not at all seriously) to 5 (very seriously). In addition, space was provided for participants to express in writing any comments or concerns they wished to share regarding their participation or the experiment itself.

Procedure

Prior to beginning data collection, this research was reviewed and approved by the James Madison IRB. Participants indicated their interest by responding to a prompt that was emailed out by faculty members, or they registered through an online database (SONA) for researchers. After making an appointment, participants were emailed one day prior to participating to arrive to court 3A at the campus gymnasium, University Recreation Center (UREC). This court was separated from court 3B by a retractable curtain that went the length of the court. Participants selected one of six potential 45 minute time slots between 9:00 AM-1:30 PM. In order to eschew delays, they were asked to arrive five minutes early to their scheduled appointment in basketball appropriate attire. A few prospective participants did not arrive in appropriate dress, so they were turned away and invited to reschedule if they desired. As they arrived, participants were provided with instructions through a typed script. To allay any concerns with the script, participants were told that the script was necessary to assure that the procedure was as similar as possible for all participants. The introduction by script lasted between 2-3 minutes, and participants were instructed to hold questions until the introduction was read in full.

To begin, the lead researcher verified that participants met the prerequisite requirement of having played organized basketball in high school. Subsequently,

participants were asked to provide informed consent and demographic information. As part of the informed consent, participants were made aware that they would be videotaped during their shooting performance and that their performance would be judged by a few basketball experts. In turn, they were assured anonymity, both in video and questionnaire responses, and were provided the opportunity to decline to participate if they were uncomfortable with these parameters. After completing the demographic form and offering consent, participants were briefed that they would be performing in a free-throw shooting competition that involved two rounds to test the effects of stress on performance.

In actuality, the competition was a ruse and was intended to test the effects of outcomes on future performance. Accordingly, the script informed participants that because it would be difficult to schedule competitors at concurrent times, they would instead be competing against an opponent who had already completed the shooting task. Participants were not made aware of their illusory opponents score until after completing their shooting round. As previously noted, participants were informed that they would be videotaped and that a performance-dependent cash incentive (up to $5) was involved. In the manner of Reeves, Tenenbaum, and Tidor (2007), participants were videotaped and provided with a cash incentive to induce stress. Further, to insure impartiality, all participants were awarded $5 upon the conclusion of their participation.

Next, participants were instructed that before their first round of shots they would be allotted 10 practice shots to warm-up and become acclimated with the performance setting and the basketball. Shots were taken on a regulation size basket (10 feet high) with either a women's basketball (28½ inch circumference) or men's basketball (29 inch

circumference) from the free-throw line 15 feet away from the basket. Participants were encouraged to give an honest effort and, though there was no time limit, they were told that they should complete their shots in a timely fashion. While the participants shot, the researcher was contemporaneously situated by the camera which was placed on the left baseline in line with the backboard. In between shots, participants retrieved their own rebounds. During both the practice and competitive rounds the researcher called out the number of shots that the participants had taken after each fifth shot. Once the practice round was complete, the researcher asked the participants to pause and verified that they understood the procedure. When the participants indicated that they understood the instructions the lead researcher resituated next to the camera and hit record. Once the camera was playing, the lead researcher instructed the participants that they could begin shooting their first round. The participants then shot 20 free throws while the researcher tallied their made baskets.

Following their first round, participants received their randomly assigned outcome (i.e., win or lose). This use of deception was critical and necessary to determine the effects of outcomes on performance evaluations and future performance. Participants were assigned to either the positive outcome group (i.e., won by one made basket) or the negative outcome group (i.e., lost by one made basket). The lead researcher asked that participants wait while they referred to the master list (a list of 52 invented scores ranging from 1-19). Prior to participating, participants were randomly assigned to an outcome (i.e., win or lose) on the basis of their experience in order to ensure that experience level was comparable for each group. However, about midway through data collection the means for the two outcome groups began to diverge–the negative outcome group

averaged more made baskets following round 1. Prior to conducting data collection it was stipulated that, within reason, the groups should be approximately even on ability level (i.e., free-throws made) following round 1. Thus, it was decided that, since the group means were diverging, a new grouping criterion would be implemented. Henceforth participants were assigned to groups following their first performance based on their ability level (i.e., baskets made) in an attempt to even the means for made baskets in round 1. Despite the change in grouping benchmark, the groups remained resemblant in terms of experience, age, and gender (see Group Demographics). Further, because their outcome could not be manipulated, it was agreed that athletes who obtained a perfect score (20 made baskets) or who did not convert a single free throw were to be allowed to continue with the experiment, but their results were to be ultimately eliminated from the data set. In reality, no participants missed or made all 20 shots.

After they received their arbitrary outcome the participants were asked to complete the CSAI-2, the self-perceived performance item, and the post-performance improvement item. While responding, participants were asked to reflect on their first round performance. Before the second round began, the lead researcher asked that the participants double-check to make sure that all questions had been answered. Once the participants verified that each item was completed, they stepped up to the free throw line again. Participants were told that they had been randomly assigned to a new competitor and that they should once more provide maximal effort. As before, the researcher hit record and tallied the number of shots made during the second round of competition while standing by the camera.

Once participants concluded their second round of shooting they were asked to fill out a brief exit survey. When the exit survey was complete they were awarded the $5 payout. Additionally, they were provided with a typed debrief of the study, and the lead researcher explicitly detailed where contact information for both their rights as a participant and questions about the research itself was located (i.e., on the back). Finally, the researcher politely asked the participants not to speak to anyone about the experiment and they were subsequently thanked for participating. In total, each appointment typically lasted between 20-25 minutes.

Because the researcher was concerned that such a small pool of potential participants (i.e., basketball players in the participant pool) may discuss the procedure and purpose of the study with future participants, the researcher opted to delay a full debrief on the use of deception until a later date. Once all of the data had been collected and stored, the researcher made participants aware of the use of deception and its purpose in a follow-up debriefing through email. In this email, participants were invited to contact the researcher if they had any lingering questions or concerns.

Hypotheses

This study explored the relationship between outcomes and performance judgments. Explicitly, the study sought to determine how an outcome (e.g., winning or losing) would influence subsequent evaluations of self-perceived performance, self-perceived cognitive state anxiety, and self-perceived self-confidence when participants were asked to recall their performance. Furthermore, this study sought to determine how an outcome would affect future performance. There were four main hypotheses:

1. As suggested by the results obtained by Sacchi and Cherrubini (2004) with physicians and Jones et al. (1997) with farcical scholarship applications, it was hypothesized that the positive outcome group would have more favorable perceptions of their performance when compared to the negative outcome group.

2. Furthermore, as evidenced by Woodman and Hardy's (2003) meta-analysis, it was hypothesized that the positive outcome group would experience less perceived cognitive state anxiety than the negative outcome group.

3. Additionally, in accordance with the results of Jones et al. (1993) study of teenage gymnasts, it was hypothesized that the positive outcome group would report greater self-confidence than the negative outcome group.

4. Finally, as intimated by the dynamometer results of Hutchinson et al. (2008), it was hypothesized that the mean difference between the scores in Round 1 and Round 2 for the positive outcome group would be significantly different than the mean difference for the negative outcome group. Moreover, in round 2 the positive outcome group would make more free throws than they did in round 1 and the negative outcome group would make fewer free throws in round 2 than they did in round 1.

Results

Group Demographics

The proposed hypotheses were explored through SPSS. To begin, the exit survey results, Round 1 performance, and the demographics for the participants in the negative outcome and positive outcome groups were compared to establish how similar these two groups were. The positive outcome group (n = 31, M_{age} = 19.4 years, SD = .84 years) was comprised of 26 males and 5 females. This group included 17 freshmen, 10 sophomores,

2 juniors and 2 seniors. Further, 11 of these participants marked recreational basketball as being their highest level played, while 4 listed freshmen basketball, 9 listed junior varsity, 6 listed varsity, and 1 listed club basketball at the university level as their highest level achieved. For the negative outcome group (n = 29, M_{age} = 19.7 years, SD = 1.2 years) there were 23 males and 6 females. This group included 15 freshmen, 6 sophomores, 3 juniors, 4 seniors, and 1 graduate student. Further, 15 of these participants listed recreational basketball as their highest level achieved, while 1 listed freshmen basketball, 4 listed junior varsity, 6 listed varsity basketball, and 1 listed club basketball at the university level as their highest level achieved. The two groups did not differ in performance (i.e., baskets made) following the first shooting trial (t(58) = -.153, p = .879, *95% CI* [-2.35, 2.02]; positive outcome group, M = 9.42, SD = 3.52; negative outcome group, M = 9.58, SD = 4.77). Further, both groups derived a similar amount of enjoyment from the experiment (positive outcome group, M = 4.39, SD = .67; negative outcome group, M = 4.38, SD = .78) and indicated that they took the task seriously (positive outcome group, M = 4.35, SD = .75; negative outcome group, M = 4.59, SD = .63). As these numbers suggest, both the negative outcome and positive outcome groups were rather similar in makeup.

The Effects of Outcomes

To explore the four major hypotheses, a series of t-tests was conducted. It was hypothesized that the positive outcome group would reflect on their performance more favorably than the negative outcome group. Per the results of the self-perceived performance item, the first hypothesis was not supported (t(58) = -1.746, p = .086, 95% CI [-1.266, .087]; d = .45). Though this difference was not statistically significant, a

Cohen's *d* test suggested that outcome had a medium effect on performance evaluation (see Cohen, 1988, for effect size interpretations). Contra the hypothesis, the negative outcome group ($M = 4.14$, $SD = 1.51$) rated their performance as stronger when compared to the positive outcome group ($M = 3.54$, $SD = 1.09$). In general, the negative outcome group rated their performance as being slightly strong, while the positive outcome group rated their performance as being slightly poor. This result suggests that performance evaluations may be related (see Cohen's *d*) to assigned outcomes such that poor results produce stronger evaluations of performance. This finding is peculiar and it goes against the hypothesized findings.

Regarding hypothesis 2, it was proposed that the positive outcome group would experience less perceived cognitive anxiety as detailed by the CSAI-2. However, no significant difference was observed between the two groups' perceived cognitive anxiety scores ($t(58) = 1.88$, $p = .066$, 95% CI [.-5.23, .14]; $d = .49$). Though this result was not statistically significant, a Cohen's *d* test suggested that outcome had a medium effect on cognitive anxiety in the hypothesized direction. In general, the positive outcome group ($M = 17.87$, $SD = 4.43$) experienced lower levels of cognitive anxiety compared to outcome group ($M = 20.41$, $SD = 5.9$) while both groups experienced low cognitive anxiety overall. This result suggests that perceptions of cognitive anxiety were affected (see Cohen's *d*) by assigned outcomes such that poor results increase perceptions of cognitive anxiety and vice versa. This finding is notable because it is in accord with the original hypothesis (i.e., hypothesis 2).

In addition, it was hypothesized that the positive outcome group would experience greater self-confidence following Round 1 as detailed by the CSAI-2. However, no

significant differences were revealed between the two groups' self-confidence scores ($t(58) = .771$, $p = .444$, 95% CI [-,1.87, 4.22]; $d = .2$). In general, the positive outcome group ($M = 26.96$, $SD = 4.73$) and the negative outcome group ($M = 25.79$, $SD = 6.93$) experienced similar levels of self-confidence. This result suggests that on average both groups experienced high self-confidence when completing the shooting competition. As discussed later, self-confidence appraisals appeared to be more dependent on Round 1 performance rather than the assigned outcome.

Finally, it was hypothesized that the mean difference scores between the shots made in Round 1 of the competition and Round 2 of the competition would be significantly different for the two groups. More specifically, it was believed that the positive outcome group would, on average, make more shots than they did in the first round and the negative outcome group would make fewer shots than they did in the first round. The divergence of the mean difference scores between rounds was then compared. However, no significant difference between the mean differences scores for the positive outcome group ($M_{diff} = .68$, $SD = 3.02$) and the negative outcome group ($M_{diff} = -.28$, $SD = 3.61$) was observed, $t(58) = ..486$, $p = .642$, 95% CI [-1.32, 2.12]; $d = .12$. These results suggest that round 2 (i.e., future performance) was not affected by the assigned outcome. This hypothesis is explored further in the post-hoc analyses.

Post-hoc Analyses

Relationship between self-confidence, cognitive anxiety, performance evaluations, and performance. In an attempt to better understand the findings of the current study, to explore post-hoc inquiries, and to examine potential moderators, the data was further scrutinized. First, a series of correlations was investigated. Regardless of

group, self-confidence was positively related to performance (Round 1, $r = .54$, $p > .001$; Round 2, $r = .39$, $p = .002$). As is expected, this suggests that athletes felt more confident when they made more baskets. Conversely, negative correlations between cognitive anxiety and performance were observed, although the relationships were not significant (Round 1, $r = -.2$, $p = .12$; Round 2, $r = -.17$, $p = .19$). As such, the self-confidence and cognitive anxiety scores were negatively related to each other ($r = -.46$, $p < .001$). Further, performance ratings were positively related to actual performance (Round 1, $r = .77$, $p < .001$; Round 2, $r = .42$, $p = .001$), which advises that the athletes' performance evaluations were not strongly affected based on the result. Accordingly, performance in Round 1 and Round 2 was strongly related ($r = .68$, $p < .001$), and self-confidence scores were also strongly related to performance evaluations ($r = .48$, $p < .001$).

Gender discrepancies regarding future performance. Despite the non-significant result, the mean difference hypothesis (i.e., hypothesis 4) deserves further detail. When females ($n = 11$) and one outlier[1] from the negative outcome group were removed from the analysis, a *t*-test of this abridged dataset ($n = 48$; positive group $n = 26$; negative group $n = 22$) revealed results that deserved mention ($t(46) = 1.88$, p = .067, 95% CI [3.42, - 0.12]; $d = .55$). While the result is not statistically significant, a Cohen's *d* test indicates that assigned outcome had a medium effect on second round performance. In fact, on average the positive outcome group (Round 1 $M = 9.15$, $SD = 3.66$; Round 2 $M = 10.08$, $SD = 3.77$) made nearly one more free throw during their second round of competition. Conversely, on average the negative outcome group (Round 1 $M = 10.41$, $SD = 4.64$; Round 2 $M = 9.68$, $SD = 4.92$) made nearly one fewer free throw during their second round of competition. This result suggests that outcomes may have a greater

effect on males when it comes to future performance[2]. Specifically, positive results may lead to stronger performances in a subsequent trial and negative results may lead to weaker performances in a subsequent trial. Interestingly, Woodman and Hardy's (2003) meta-analysis of self-confidence, cognitive anxiety, and performance found significant results for males, but not females. All this recommends that gender may play a meaningful role in the dynamic between outcomes, future performance, and feeling states. Additional investigation of this hypothesis can be found in the discussion section.

Skill level effects. Following the statistical exploration of the hypotheses and the sequence of correlations, it was determined that the skill level variable (denoted by the highest level of organized basketball achieved) deserved further attention to determine whether outcomes differentially affect varying levels of experience. There were five levels of skill: recreational (n = 26), freshman (n = 5), junior varsity (n = 13), varsity (n = 14), and beyond varsity (n = 2). Keeping in mind sample size concerns, there are several means that bear mention[3]. Given that only means are reported, it must be stated that no casual inferences can be drawn from these numbers. Nevertheless, these results are provided as evidence that future research should further scrutinize the relationship between skill level and outcome effects. This variable may have acted as a moderator impeding the ability to isolate an outcome bias and in turn forcing the hypotheses to not be supported. Because just five participants listed freshman basketball as their highest level achieved and only two participants listed beyond club as their highest level achieved, neither of these groups are included in the forthcoming discussion.

Per their mean difference scores (see Table 1), recreational and junior varsity athletes improved by nearly half a shot in round 2, while varsity athletes improved by a

full shot. As would be expected, varsity athletes made, on average, more shots (about 2 more shots) than junior varsity and recreational level athletes. Finally, when only males (n = 48) are included in the comparison, a few mean differences become more pronounced or entirely new relationships emerge[4]. For instance, recreational athletes do not improve in their round 2 performance, but junior varsity and varsity athletes do improve. These means intimate that skill levels are unique and that outcomes may distinctly affect different skill levels. For further detail on skill levels effects see Table 1.

Table 1

The Effects of Skill on Feeling States, Performance, and Performance Evaluation

	Recreational (n = 26)		Freshman (n = 5)		Junior Varsity (n = 13)		Varsity (n = 14)		Beyond Varsity (n =2)	
	Win (n=11)	Lose (n=15)	Win (n=4)	Lose (n=1)	Win (n=9)	Lose (n=4)	Win (n=6)	Lose (n=8)	Win (n=1)	Lose (n=1)
Variable	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Cognitive Anxiety	18.2 (4.8)	20.1 (5.8)	17.3 (3.2)	24 (0)	17.6 (5.4)	19.8 (5.4)	18.3 (4.1)	20.3 (7.3)	17 (0)	25 (0)
CA Total	19.3 (5.4)		18.6 (4.1)		18.2 (5.3)		19.4 (6)		21 (5.7)	
Self-Confidence	27.5 (5.7)	24.1 (6.7)	27.5 (3.5)	23 (0)	26.4 (4.6)	26.3 (4.2)	25.8 (5)	29.8 (8)	30 (0)	20 (0)
SC Total	25.6 (6.4)		26.6 (3.6)		26.4 (5.3)		28.1 (7)		25 (7.1)	
Performance Evaluation	3.5 (1.2)	4.0 (1.5)	4 (1.2)	5 (0)	3.4 (1.1)	3.8 (1.5)	3.3 (1.0)	4.5 (1.5)	4 (0)	4 (0)
PE Total	3.8 (1.5)		4.2 (1.1)		3.5 (1.2)		4 (1.4)		4 (0)	
Round 1	8.3 (3.8)	8.8 (4.7)	11.5 (3.7)	3 (0)	10 (2.5)	7.8 (4.0)	8.7 (4.0)	12.4 (4.5)	13 (0)	13 (0)

Round 1 Total	8.6 (4.3)		9.8 (5)		9.3 (3.1)		10.8 (4.5)		13 (0)	
Round 2	9.7	8.6	9.7	6	10.1	8.5	10.3	13	14	13
	(4.3)	(4.5	(1.3)	(0)	(3.4)	(5.1)	(4.5)	(3.1)	(0)	(0)
Round 2 Total	9.1 (4.3)		9 (2)		9.6 (3.9)		11.9 (3.8)		13.5 (0.7)	
M_{Diff} Round 1 to Round 2	1.5	- 0.2	- 1.8	3	0.1	0.8	1.7	0.6	1	0
	(3.5)	(4.0)	(2.6)	(0)	(2.2)	(4.3)	(3.3)	(3.3)	(0)	(0)
M_{Diff} *Total*	0.5 (3.8)		- 0.8 (3.1)		0.3 (2.8)		1.1 (3.2)		0.5 (0.7)	

Note. Cognitive anxiety (CA) and self-confidence (SC) scores are derived from the CSAI-2 (Martens, Vealey, & Burton, 1990); scores range from 9-36. The personal performance evaluation (PE) item is adapted from Sacchi and Cherubini (2003); scores range from 1-7. Round 1 and Round 2 means are represent made baskets (out of 20 attempts). *Total* scores represent the means and standard deviations for skill level regardless of outcome (i.e., across groups).

Discussion

Summary of Results

This research examined the effects of outcomes in sport on evaluations of performance and feeling states. Specifically, participants took part in a free-throw shooting competition wherein their illusory opponent had already completed his or her shooting round. Following Round 1, participants were provided with a contrived result (i.e., won by one or lost by one made basket) to determine whether results would affect their evaluations of their performance, their appraisals of their self-confidence and cognitive anxiety, and their performance in a second round.

Outcomes and self-perceptions of performance. It was hypothesized that the favorable outcome group would rate their performance as better when compared to the

negative outcome group. However, the two groups did not significantly differ in this manner. In fact, the negative outcome group rated their performance as better, though this result was not statistically significant. This finding sits in stark contrast to much of the extant outcome literature. For example, Italian physicians rated their skill of diagnosis on the basis of its outcome (Sacchi & Cherubini, 2004).

Several possible explanations exist as to why the hypothesized relationship between outcomes and self-perceptions of performance was not seen. First, the free-throw shooting tournament may not have been appropriate to evoke the emotions (e.g., anxiety) experienced during actual competition. Greater impressions of cognitive anxiety have been related to performance, such that as cognitive anxiety increases performance tends to decrease (Woodman & Hardy, 2003). However, athletes in the current study experienced levels of cognitive anxiety (positive outcome group, $M = 17.87$, $SD = 4.43$; negative outcome, $M = 20.41$, $SD = 5.9$) that on the whole were similar to high school male (this research had mostly male participants) basketball players ($M = 20.92$, $SD = 6.11$; Martens et al., 1990) prior to competition. Likewise, participants reported levels of self-confidence (positive outcome group, $M = 26.97$, $SD = 4.73$; negative outcome group, $M = 25.79$, $SD = 6.93$) that were similar to high school male basketball players ($M = 24.73$, $SD = 5.52$) prior to competition. Furthermore, participants reported that they took the free-throw shooting task seriously. These findings indicate that the free-throw shooting competition produced feeling states tantamount to those experienced in actual competition.

Alternatively, it may be that outcomes affect the athlete less than those in other professions. The results of this study may suggest that there is a sort of buffer mechanism

in response to losing, wherein the athlete reacts by resolving that his or her performance was not that poor. This explanation appears somewhat improbable when you consider that performance in a dynamometer exercise was severely hampered following a poor outcome and vice versa (Hutchinson et al., 2008). Further, outcomes have been shown to influence how coaches and athletes interact with the game. For instance, NFL coaches tended to reshape their strategy in response to a loss or continue on with their current strategy following a win (Lefgren et al., 2011).

From our view, it is most likely that these numbers represent statistical flukes and that further data collection would bear out that the relationship is truly either null or that positive outcomes improve performance evaluations and vice versa. After all, physicians have been shown to rate their performance based on outcomes (Sacchi & Cherubini, 2003) and NBA coaches are more likely to manipulate their starting lineups following a loss (Lefgren et al., 2013). Accordingly, future research should aim to further detail the relationship between outcomes and performance evaluations.

Outcomes and self-perceptions of cognitive state anxiety and self-confidence. It further was hypothesized that the positive outcome group would appraise their cognitive state anxiety and self-confidence related to their Round 1 performance as lower and higher, respectively, when compared to the negative outcome group. However, the two groups did not differ in their cognitive state anxiety scores. This result is not entirely surprising. Indeed, regarding cognitive anxiety, Woodman and Hardy (2003) promulgated discrepant findings in their meta-analysis of the CSAI-2. For instance, in a group of teenage gymnasts, cognitive anxiety was negatively related to performance, while on the other hand, anxiety was positively related to performance for a group of

male wrestlers cognitive anxiety was positively related to performance. The current study observed that cognitive anxiety was negatively related to performance (Round 1 r = -.2; Round 2 r = -.17), which aligns well with the correlations reported in Woodman and Hardy's (2003) meta-analysis (r = -.10, r = -.22 for males). The conversation becomes more tenuous when you consider that some argue that anxiety can be interpreted as either facilitative or debilitative (Woodman & Hardy, 2003). Athletes may have high anxiety, but interpret this feeling state as being facilitative to performance. Accordingly, they may perform well under such cognitive pressure and vice versa. All this is to say that it remains unclear how performance interacts with anxiety. Thus, future research should strive to elucidate further how cognitive anxiety relates to performance and outcomes.

That being said, a Cohen's d test recommends that the athletes' perceptions of cognitive anxiety actually were swayed by assigned outcome to a medium effect (d = .49). The positive outcome group reported lower levels of cognitive anxiety when they reflected on their round 1 performance than did the negative outcome group. Thus, outcomes may have the capacity to tinker with cognitive anxiety. Furthermore, this result signals that other reflections on performance may be distorted as a product of outcomes as well. For instance, an athlete may be prone to recalling positive aspects of a performance following a win or other positive result while the opposite may be true for an athlete following a loss or a negative result. As an example, a tennis player may reason that he or she lost because they have a poor forehand shot when in actuality the performance could be chalked up to reasonable variability in forehand stroke performance. An even more deleterious possibility is that an athlete may take the result as an indication of ability and what must have happened (see Nisbett & Wilson, 1977, for a

review of faulty causal narratives) and will deny evidence to the contrary. They may reason that, "well I lost, so I must have a poor forehand stroke." In doing so, an athlete may be unable to recognize what skills they are proficient in and what skills need to be honed. There are innumerable analogues in other sports and coaches, practitioners, and other sport actors are no less susceptible to this biased reasoning. To be sure, more research is needed, but if this account is corroborated by future research this finding would undoubtedly be of interest to coaches, athletes, and practitioners alike.

In addition, the results of this study revealed that the positive and negative outcome groups did not differ in their self-confidence scores. This suggests that self-confidence was not depressed as a result of poor outcomes or enhanced by a positive outcome. Interestingly, self-confidence was strongly related to performance across both groups of participants. Hanton et al. (2004) posited that athletes with greater self-confidence tend to recall stronger performances and vice versa. Thus, it makes sense that participants who performed more strongly in Round 1 reported having higher self-confidence. There is not yet a consensus on how self-confidence relates to performance. In the current study, self-confidence was strongly related to performance (Round 1, r = .54; Round 2, r = .39). Woodman and Hardy (2003) also determined self-confidence was related to performance (r = .24) through their meta-analysis (r = .29 for males) of this relationship. Nonetheless, 10% of the studies they reviewed found a negative relationship between self-confidence and performance. In reaction, future research should seek to determine if a relationship exists between outcomes and self-confidence.

Outcomes and subsequent performance. Finally, it was hypothesized that the mean difference between the two groups between their Round 1 and Round 2 scores

would be significantly different. That is, the positive outcome group would convert more free throws from Round 1 to Round 2 and the negative outcome group would convert fewer free throws. Although this hypothesis was rejected, it must be noted that for males both of the groups mean difference scores between rounds trended in the hypothesized directions. More to the point, males in the positive outcome group made nearly one more free throw during their second round, while those in the negative outcome group made nearly one fewer free throw. As suggested by Cohen's *d*, this performance discrepancy (*d* = .55) is indicative of a medium effect. Such an effect is undoubtedly of interest to sport professionals. Further, these results line up with those found in Denes-Raj and Epstein's (1994) jellybean procedure where discouraging results (i.e., not picking a winning jellybean) led participants to endorse a suboptimal strategy in subsequent trials. Similar performance degradations have been observed in investment tasks (Ratner and Herbst, 2005), gaming strategy (Erev & Roth, 1998), strength training (Hutchinson et al., 2008) and NFL play calling (Lefgren et al., 2011). Given that performance is such a critical variable in sport, it is incumbent upon the field to invest in research and statistical analyses that will aid in creating a better understanding of the response to outcomes of coaches, athletes, and teams in general. In the current, this result signals that coaches, athletes, and teams should be wary of outcome effects and be skeptical of the notion that a given outcome – win or loss, success or failure – provides valuable input when evaluating performance.

In summation, a series of *t*-tests provided no support for the four main hypotheses. However, further scrutiny of the data yielded two statistics of note. First, regarding hypothesis 2, assigned outcome tended to affect perceptions of cognitive

anxiety to a medium effect (d = .49). Those who received a positive outcome reported lower levels of cognitive anxiety during round 1. Second, regarding hypothesis 4, round 2 (i.e., future) performance for males was also affected by assigned outcome to a medium effect (d = .55). Males who were provided with a positive outcome tended to perform better in round 2 and males who received a negative outcome tended to perform worse in round 2. These findings and their implications are explored in further detail in the subsequent discussion.

Limitations

There are limitations to the current research that deserve discussion. Specifically, there are concerns about external validity, test environment, magnitude effects, and sample size. It must be stated that acquiring external validity is always an issue when studying sport (Woodman & Hardy, 2003). Undeniably, it is difficult to replicate the pressure and anxiety associated with competition. However, it should be noted that in an attempt to augment stress, the current study used cash incentives (Arkes et al., 1986) and videotaped the participants while they were shooting (Reeves et al., 2004). Moreover, participants indicated that they experienced similar levels of self-confidence and cognitive anxiety during the task as their peers prior to competition (Martens et al., 1990). Nevertheless, down the road sport researchers are encouraged to construct test environments that mirror the athletes' performance milieu and allow for outcome manipulation.

Another concern deals with the test environment. Though the divider was down and the court was reserved, gym patrons could still use the court on the other side of the divider. Patrons would come and go, so this meant that sometimes there were others in

the gym and sometimes there were not. Noise pollution from the adjacent court may have unsettled participants while they were shooting and in turn meddled with their performance corrupting our results. Further, on a couple occasions the experiment was disrupted by patrons who wished to use the other half of the court that the participants were using. These interlopers were hurried away immediately and the disruption was deemed minimal. Nonetheless, despite best efforts to limit these interruptions, participants who performed in the presence of other patrons may have performed worse than they otherwise would have. As such, future research would ideally take place in a more intimate testing environment to isolate any potential outcome effects.

Additionally, there are concerns that the outcome manipulation (i.e., win by 1 or lose by 1) was not strong enough to observe the hypothesized effects. However, past research suggests that the magnitude of outcomes does not exacerbate the results of outcome effects (Lefgren et al., 2014). Furthermore, minimizing the magnitude of a loss or a win allowed us to retain both elite and poor performances. No doubt it is valuable to be able to retain and manipulate the outcome for an athlete who makes 19 free throws. Ostensibly, making 19 shots indicates a strong performance, but an athlete may rate their performance worse if they lose. As such, it was determined that a one shot magnitude was ideal because it permitted both possessing more participants results and studying the severity of outcome effects. Nevertheless, exploring the effects of outcomes in the presence of larger wins and losses is a worthwhile pursuit and may yield more conclusive results.

Finally, the sample size for each group ($n = 29$ for positive outcome group, $n = 31$ for negative outcome group) was small. For instance, in order to have ideal power ($1 - \beta$

= .8) to find a medium effect (d = .5) in a t-test, Cohen (1992) recommends a sample of 64 per group. Had more participants been included, statistically significant results may have been obtained (see specifically hypothesis 2 on cognitive anxiety and, for males only, hypothesis 4 on subsequent performance). However, the number of prospective participants that met the prerequisite criteria of having played organized basketball in high school in the participant pool was limited. Indeed, even obtaining the suggested sample size was originally a concern. Thus, future replications of the current study, or future research of outcome bias in the broad, should seek to employ a more appropriate number of participants.

Future Research

Beyond the lines of future research implied in the afore discussion of the hypotheses, several other variables of import and their potential moderating effects are worthy of note. It is recommended that these variables and their relations to outcomes be explored. For instance, the current study does not appropriately address the potential issue of gender. The results for the current study's small sample of females suggest that outcomes may not have a strong impact on performance. Moreover, the performance for females in the positive outcome group did not improve and the performance for females in the negative outcome group did not suffer. Conversely, for males, performance improved when they were provided with a positive outcome and decreased when they were provided with a negative outcome. However, the sample size (n = 11) for females was wanting, and gender remains a variable of interest. Thus, the interaction between sex and outcome effects should be further explored in future research.

Another matter that is not addressed properly in the current research is duration of impact. If outcome effects are fleeting, they may not be of great import to most athletes and coaches because in most sports the gap between competitions is several days. However, there is evidence to suggest that outcome effects remain prominent for an extended period of time. As a result, athletes and coaches who are associated with an outcome (e.g., a strong performance, winning a championship, or a losing season) may be assigned distorted performance evaluations (e.g., too much blame or too much praise) in subsequent competition as a product of this association. For example, a group of gymnastic judges were shown a tape of a gymnast who either had a perfect routine or a routine with an error (Ste-Marie & Valiquette, 1996). One week later, the judges were asked to grade another performance (both groups viewed the same routine) from the same gymnast. Judges who viewed the perfect routine graded the subsequent performance more favorably. To test whether or not outcome effects are fleeting and whether the aforesaid findings will generalize, future research should explore further the lasting duration of outcome effects.

In addition, different sports carry with unique stressors. Thus, you may find that long distance runners (Martin & Gill, 1991) perform better in the presence of high self-confidence and low cognitive anxiety while the inverse is true of netball players (Edwards & Hardy, 1996). Outcomes may work similarly, and they may have a greater impact in certain sports. On the contrary, type of sport may have little impact while individual differences may be more pertinent. Some athletes, regardless of sport, may respond to results to a greater extent and place more emphasis on them when evaluating

performance and feeling states. Determining where outcome effects fall on the sport versus individual continuum should be parsed in future research.

Finally, as it relates to outcome bias, skill level has the potential to be a meaningful mediating factor. Indeed, the current research observed a potential impact of skill level. Varsity level athletes tended to exhibit greater self-confidence, made more baskets, rated their performance as better and improved more from Round 1 to Round 2 than any other skill level. How might elite level athletes, where the pressure to succeed is exacerbated, respond to outcomes? Jones, Hanton, and Swain (1994) suggested that although impressions of anxiety may not shift dramatically as athletes progress through the ranks, feelings of self-confidence do. Namely, that as athletes begin to compete at more elite levels and acquire more skill their self-confidence is augmented. Consequently, sport researchers should move to determine to what degree skill level differentially works to mitigate or intensify outcome effects.

Outcome Bias and the Belief in the Hot Hand and Clutch Players

Gymnastics judges have been shown to attach certain outcomes (e.g., a routine with no mistakes) to gymnasts and to judge future routines on the basis of these attached outcomes (Ste-Marie & Valiquette, 1996). Similarly, it is not uncommon for NBA players to be attached to specific outcomes (e.g., making an important shot late in a game). These ephemeral performances inspire grand exaltation and lead the media and fans to respond in kind by bestowing these players with gaudy appellations. One example was the moniker given to the Los Angeles Lakers' Kobe Bryant, "The Black Mamba", an ode to the large, extremely aggressive and venomous snake found in sub-Saharan Africa. Bryant received this label because he was viewed as a peerless talent in crucial, late game

scenarios (Rovell, 2016). Another example is Vinnie "Microwave" Johnson who spent most of his career with the Detroit Pistons. Johnson's apparent streak shooting ability inspired work (Larkey, Smith, & Kadane, 1989) that attempted to trample on research documenting the hot hand fallacy (Gilovich, Vallone, & Tversky, 1985) or the fallacious belief in streaks. Ultimately, the conclusions of Larkey et al. (1989) were debunked when it was found that the statistics that were used were invented (Tversky & Gilovich, 1989). Going far beyond just basketball, this fallacy has been documented in many more sports including golf (Clark, 2005), and baseball (Albert, 1993), among others. In fact, this deception is so ubiquitous that it's also been observed in disciplines outside of sport, perhaps most notably in stocks and mutual fund performance (Kahneman, 2011; Sirri & Tufano, 1998).

These celebrated players are in turn counted on later to perform in deciding moments of competition; this class of athletes is commonly said to be clutch (Solomonov, Avugos, & Bar-Eli, 2015). However, Solomonov et al. (2015) reveal that this perception of clutch players is arguably a misnomer. Players who were denoted as clutch by NBA experts performed no better in pressure situations than their ostensibly ordinary peers. Instead, they experienced typical performance degradations in high-pressure scenarios (Cao, Price, & Stone, 2011). Nonetheless, the belief in the hot hand and the notion of clutch players is pervasive and many sport professionals and some sport researchers (e.g., Miller & Sanjurjo, 2014) would scoff at the above research. This belief has so much sway that teammates and coaches tend to orient the ball to shooters with the perceived hot hand or to clutch players during critical moments of a game (Attali, 2013; Csapo, Avugos, Raab, & Bar-Eli, 2015). In this way they feed the myth, because clutch players and

players with the perceived hot hand get more shots and thus are provided with more opportunities to succeed in pressure situations. The fact that these players are often no more successful in these situations does not appear to bear on the minds of most sport professionals.

Though the debate continues, in the least, the relationship between belief in ability (e.g., how clutch a player is) and actual performance is far weaker than many purport it to be (Moskowitz & Wertheim, 2011; Vergin, 2000). Thus, coaches and athletes should be wary of their perceptions of players' abilities in crucial moments. If they are not, they may endorse and utilize a deficient strategy in these critical moments and settle for an uncommonly difficult shot when other options exist. Here, coaches and athletes fall prey to an outcome bias because they tend to, in lieu of the failures, recall moments of success. For perspective, Kobe Bryant may have made many significant baskets in his career, but this is a product of his taking many significant shots. He missed a great deal as well and, in fact, his performance in crucial situations is opposite of what his moniker suggests–he performs poorly (Abbott, 2011). Perhaps the Lakers would have won more games had they entertained other options during pivotal moments.

Explaining Outcome Effects: The Availability Heuristic, Confirmation Bias, the Narrative Fallacy, and Randomness

An obvious rejoinder to the suggestion of an outcome bias is, how did this bias come to be? Though this was not the immediate purpose of this review and research, a cursory explanation is offered now so that future research and discussion may be steered in the right direction. Here, three phenomena bear mention: the availability heuristic (Tverksky & Kahneman, 1973; Schwarz, Bless, Strack, Klumpp, Rittenauer-Schatka, &

Simons, 1991), the confirmation bias (Nickerson, 1998), and the narrative fallacy (Nisbett & Wilson, 1997; Taleb, 2007; Wilson, 2002).

According to the availability heuristic, when considering decisions or topics, we do not pull all relevant information to the fore (Kahneman & Tversky, 1973). Instead, we rely on a few specific examples that are easily called to mind. In addition, Schwarz et al. (1991) determined that ease of recall is also an influential factor when considering decisions or topics. It follows that it is easier to recall two examples of an event than it is to recall three examples of an event and so on. The more examples we are requested to recall, the less confident we are in our conclusions. When left to our own devices, however, we tend to take an example or two and then make a judgment. In turn, we are prone to notice instances that align with our original judgments. That is to say that we succumb to a confirmation bias (Nickerson, 1998). Thus, a coach may recognize when a player who has been deemed clutch makes a basket in a crucial situation, but fail to acknowledge a disconfirming instance where this player misses a basket at a crucial moment (recall the Kobe Bryant anecdote, Abbott, 2011). In the current study, participants in the negative outcome group may have recalled losing (i.e., performing poorly) in round 1 and subsequently performed worse than they otherwise would have.

The narrative fallacy speaks to our predilection for shaping narratives around facts (Taleb, 2007). It is not enough for us to know that an event occurred; in tow we require an explanation for said event. These explanations put us at ease even when they are unfounded or erroneous (Taleb, 2007; Nisbett & Wilson, 1977). In an extreme example, research subjects were asked to watch a documentary about the plight of Jewish indigents and to afterwards rate their reactions (e.g., how sympathetic they found the

main character to be) to the film (Nisbett & Wilson, 1977). In one condition, participants were subjugated to distracting noise (i.e., a power saw) directly outside of the viewing room. The noise condition reported that the noise had affected their reviews when in truth their ratings were no different than the control group. So engrained is this need for an explanation that these subjects mistakenly reported that their ratings had been affected.

Further complicating this discussion, many do not have a firm grip on randomness and chance events (Burns & Corpus, 2004). For instance, a great deal believe that a random-number generator is more likely to spit out a variegated sequence of digits such as 4891, than a repetitive sequence such as 2222, even though both series are just as likely (Chapanis, 1953). Another example of our ineptitude is seen when many do not recognize that in a progression of 200 coin flips it is likely that you will at some point observe a string of 6 or 7 heads or tails in a row (Tversky & Kahneman, 1971). Moreover, we tend to believe that truly random sequences or events do not cluster leading us to conclude that clusters are indicative of an observable pattern. This phenomenon, which has been documented in sport (Gilovich, 1984), is known as the clustering illusion (Gilovich, 1991).

In the present research, participants in the negative outcome group may have taken their result as evidence that they were more anxious than they truly were. They may have reasoned that anxiety is associated with poor performances so they must have been anxious. All this is not fully conscious reasoning, but rather the product of a lifetime of reinforced rationalizing (Wilson, 2002). We do not always know why an event took place, but we prefer to believe that we do (Taleb, 2007; Wilson, 2002).

This fallacy is of particular consequence when considering outcomes that are unexpected. When the actual outcome of an event is not what was expected, this does not necessarily suggest that our a priori expectations were invalid. Indeed, when dealing with complex environs as capricious as a sporting event, one must respect the potential for innumerable outcomes. Hence, when a team experiences an unexpected loss or an athlete performs worse than expected, the outcome it is not necessarily cause for concern. It may simply be a development in line with reasonable variability. Nevertheless, it appears that it is hard for us to intuit that expected outcomes and actual outcomes will not always align.

Significance of the Study

In the main, the hypotheses of this research were not supported. Nonetheless, there are still several relevant implications for both practitioners and scholars in the field of sport and athletics to consider. The current study was unique from previous research in that outcomes were manipulated in a sport setting to determine if they would alter future performance. Thus, this research fills a gap in the extant literature because it is an empirical test of the effects of outcomes in sport. Of most interest, the current study observed an outcome bias in sport when male participants in the negative outcome group made fewer free throws (nearly one fewer) in their Round 2 performance while those in the positive outcome group made more free throws (nearly one more) in their Round 2 performance. Further, as suggested by a Cohen's *d* test, cognitive anxiety was affected by assigned outcome to a medium effect. Thus, those who were assigned a positive outcome reported feeling less anxious when they reflected on their round 1 performance.

Seemingly, outcomes have the ability to mangle our recollections and influence future performance.

As suggested earlier, there are several lines of potential research that could expand our understanding of this phenomenon. For example, will males and females differ in regards to their response to outcomes? In the current study, future performance for males and females was differentially affected by assigned outcome. Furthermore, determining what variables (e.g., mental toughness, interpretations of stress and anxiety, self-efficacy) delineate between individuals who are highly and lowly affected by outcomes is a worthwhile pursuit of future research. For instance, Jones et al. (1994) found that elite athletes tended to have more facilitative interpretations of anxiety than their non-elite peers. It is important to note that, while the impacts of potential mediating variables (e.g., self-efficacy) are worthy of interest, the purpose of the current study was to determine simply whether or not an outcome bias exists in sport. This research advances that this concern is founded, as evidenced by the impact of outcomes on subsequent performance and perceived cognitive anxiety, thus future research should strive to discern what variables act to manipulate this effect. Once the impact of these variables is better understood and outcome effects in sport are finely detailed, researchers must move to determine what can be done to mollify the effects of outcomes.

In the interim, relevant parties should not delay action. Going far beyond just free throw shooting, an outcome bias has the potential to perniciously affect performance in a variety of ways (recall the discussion in Outcome Effects in Sport). Following the preponderance of evidence that has been offered, organizations, athletes, and coaches should take steps to adumbrate plans that will counter potential ill steps brought on by an

outcome bias. Additionally, it is counseled that those in sport be cautious with the narratives they shape around results. In short, every loss or poor outcome is not necessarily the fruit of defective strategy. Nor is every win or optimal outcome always the crop of choice strategy (Gilovich, 1991; Taleb, 2007; Wilson, 2002).

Pressing forward, it is advised that practitioners broach the topic of outcome effects with the athletes and teams with whom they work. The outcome bias corpus suggests that both evaluations of personal performance (e.g., Sacchi & Cherrubini, 2008) and others' performance (e.g., Hutchinson et al., 2008) are slanted as a product of results (e.g., we lost so we played poorly). This research demonstrated that cognitive anxiety might be augmented in the face of a poor result. Additionally, future performance has been shown to deteriorate following poor outcomes (e.g., Hutchinson et al., 2008). Likewise, the current study found a trend existed for male athletes who were provided with a positive result following a round of free throw shooting. These participants performed stronger in a subsequent round of free throw shooting, while athletes who received a negative result performed worse. This is concerning because athletes and coaches are not in complete control of results—a stroke of good fortune can have fortuitous effects on outcomes while a fluke can have pernicious effects.

Furthermore, coaches should be wary of the potential impact of specific performances (i.e., an outcome bias) and in response should not overly rely on or discredit their athletes' abilities in certain situations (recall the discussion in Outcome Bias and the Belief in the Hot Hand and Clutch Players). As an example, a coach who witnesses a player performing poorly in a crucial situation should not unquestionably reason that this athlete is incapable. Instead, he or she should compare this performance

to an average performance (e.g., shooting percentage) and determine if the approach and the athlete's response was fitting. This advice holds firm for all relevant actors (e.g., athletes, practitioners, fans) in the sport domain. Likewise, coaches can be inappropriately associated with outcomes. For instance, a coach who outperforms expectations (i.e., beats the spread), but still incurs a bad result, may not be appropriately commended for his or her leadership. Athletes, management, fans, and the media should take a more objective approach when judging the work of coaches rather than just following the spurious implications of the outcome–performance relationship. Thus, practitioners should provide these groups with means to judge their performance aptly, irrespective of results.

In general, it is hoped that the findings of this research and the broader literature that was proffered will spark a conversation among sport professionals about the role that outcomes play in performance evaluations. Indeed, outcomes have been shown to have drastic effects on people's thoughts, cognitions, and actions; sport is likely no different. The ultimate aim in athletics is to increase the likelihood of success in the future. If we allow past outcomes to obfuscate this fact, we have failed to achieve this overarching goal and, worse yet, we have set athletes and coaches up for failure in the future.

Although the impetus for the misguided raillery of the Seahawks' final play call should now be clearer, the predominant question moving forward is whether or not actors in the sport domain will push to quell this untoward approach. Those who intend to heed this call should be cautioned that putative experts are liable to erroneously endorse a form of hindsight sophistry. Indeed, football minds, such as Emmitt Smith, used the result of Russell Wilson's pass as prima facie evidence that the play call was inappropriate.

Unfortunately, there is no indication that the sport intelligentsia will be disillusioned of their fascination with, and their specious confidence in, the implications of results any time soon, but it is hoped that a few visionaries will act as lodestars for greater objectivity.

Our inability to ponder possibilities and reason probabilistically leads us to unjustly lambaste players, coaches, and organizations. The conjecturing that follows poor outcomes often plays a hand in the ultimate termination of the coaches, players, and management involved. Not just in sport, but also in a broad range of fields, outcomes have deleterious impacts. Indeed, shades of a bias run rampant in newspapers, history books, performance evaluations, et cetera. Highlighting our opacity, undue plaudits are granted to those who find success. What's worse is our inexpert use of results leads us to wrongly condemn those who fail. Hard work notwithstanding, success is to some degree a product of chance–failure is too.

In our jobs and other daily activities, our blind belief in fate has potent ramifications. Unwittingly and constantly, outcomes warp our view of the world around us. However, outcomes are not preordained and we cannot allow outcomes and the backfit narratives that accompany them to dictate our understanding of events. The environment we operate in is intractably complex and we must accept that the world is full of variability and volatility. The outcomes we receive and experience are not always expected and, through no fault of our own, are often unpredictable.

References

Abbott, H. (2011, January 29). The truth about Kobe Bryant in crunch time. *ESPN*. Retrieved from http://www.espn.com/blog/truehoop/post/_/id/24200/the-truth-about-kobe-bryant-in-crunch-time

Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes*, *37*(1), 93-110.

Albert, J. (1993). A statistical analysis of hitting streaks in baseball: Comment. *Journal of the American Statistical Association*, *88*(424), 1184-1188.

Arruda, A. F., Aoki, M. S., Freitas, C. G., Drago, G., Oliveira, R., Crewther, B. T., & Moreira, A. (2014). Influence of competition playing venue on the hormonal responses, state anxiety and perception of effort in elite basketball athletes. *Physiology & Behavior*, *130*, 1-5.

Attali, Y. (2013). Perceived hotness affects behavior of basketball players and coaches. Psychological Science, 24, 1151–1156.

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, *54*(4), 569-579.

Berger, J., & Pope, D. (2011). Can losing lead to winning?. *Management Science*, *57*(5), 817-827.

Blount, T. (2015, February 2). Seattle O-boss: 'conscious' of clock. *ESPN.* Retrieved from http://espn.go.com/nfl/playoffs/2014/story/_/id/12266535/super-bowl-xlix-darrell-bevell-seattle-seahawks-play-call-made-kill-clock

Burns, B. D., & Corpus, B. (2004). Randomness and inductions from streaks: "Gambler's fallacy" versus "hot hand". *Psychonomic Bulletin & Review*, *11*(1), 179-184.

Cao, Z., Price, J., & Stone, D. F. (2011). Performance under pressure in the NBA. *Journal of Sports Economics*, *12*(3), 231-252.

Caplan, A., & Igel, L. (2015, February 2). Data, guts, and the right way to think about the Seahawks' decision to make the wrong call. *Forbes.* Retrieved from http://www.forbes.com/sites/leeigel/2015/02/02/data-guts-and-a-right-way-to-think-about-the-seattle-seahawks-decision-to-make-the-wrong-call/

Chapanis, A. (1953). Random-number guessing behavior. *American Psychologist*, *8*, 332.

Chiviacowsky, S., & Wulf, G. (2007). Feedback after good trials enhances learning. *Research Quarterly for Exercise and Sport*, *78*(2), 40-47.

Chiviacowsky, S., Wulf, G., Wally, R., & Borges, T. (2009). Knowledge of results after good trials enhances learning in older adults. *Research Quarterly for Exercise and Sport*, *80*(3), 663-668.

Clark, R. D. (2005). An examination of the "hot hand" in professional golfers. *Perceptual and motor skills*, *101*(3), 935-942.

Cohen, J. (1988). *Statistical power analysis for behavioral sciences* (2nd ed.). Hillsdale, NJ: L. Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159.

Csapo, P., Avugos, S., Raab, M., & Bar-Eli, M. (2015). The effect of perceived streakiness on the shot-taking behaviour of basketball players. *European journal of sport science*, *15*(7), 647-654.

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: when people behave against their better judgment. *Journal of Personality and Social Psychology*, *66*(5), 819-829.

Drust, B. (2010). Performance analysis research: Meeting the challenge. *Journal of Sport Sciences. 28(9),* 921-922.

Edwards, T., & Hardy, L. (1996). The interactive effects of intensity and direction of cognitive and somatic anxiety and self-confidence upon performance. *Journal of Sport and Exercise Psychology*, *18*, 296-312.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 848-881.

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(3), 288-299.

Fischhoff, B., & Beyth, R. (1975). I knew it would happen: Remembered probabilities of once—future things. *Organizational Behavior and Human Performance*, *13*(1), 1-16.

Gilovich, T. (1984). "Judgmental biases in the world of sports. In W.F. Straub and J.M. Williams (eds.), *Cognitive Sport Psychology* (pp. 31-41). Lansing, NY: Sport Science Associates.

Gilovich, T. (1991). Cognitive determinants of questionable beliefs. *How we know what isn't so: The fallibility of human reason in everyday life* (pp. 9-74). New York, NY: The Free Press.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive psychology*, *17*(3), 295-314.

Gino, F., Moore, D. A., & Bazerman, M. H. (2009). No harm, no foul: The outcome bias in ethical judgments (Working Paper No. 08-080). Retrieved from Harvard Business School NOM website: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1099464

Hanton, S., Mellalieu, S. D., & Hall, R. (2004). Self-confidence and anxiety interpretation: A qualitative investigation. *Psychology of Sport and Exercise*, *5*(4), 477-495.

Hays, K., Thomas, O., Maynard, I., & Bawden, M. (2009). The role of confidence in world-class sport performance. *Journal of Sports Sciences*, *27*(11), 1185-1199.

Hutchinson, J. C., Sherman, T., Martinovic, N., & Tenenbaum, G. (2008). The effect of manipulated self-efficacy on perceived and sustained effort. *Journal of Applied Sport Psychology*, *20*(4), 457-472.

Jones, G., Hanton, S., & Swain, A. (1994). Intensity and interpretation of anxiety symptoms in elite and non-elite sports performers. *Personality and Individual Differences*, *17*(5), 657-663.

Jones, G., Swain, A., & Hardy, L. (1993). Intensity and direction dimensions of competitive state anxiety and relationships with performance. *Journal of Sports Sciences*, *11*(6), 525-532.

Jones, S. K., Yurak, T. J., & Frisch, D. (1997). The effect of outcome information on the evaluation and recall of individuals' own decisions. *Organizational Behavior and Human Decision Processes*, *71*(1), 95-120.

Kahneman, D. (2011). Part 3: Overconfidence. *Thinking, fast and slow* (pp. 199-265). New York, NY: Farrar, Straus, and Giroux.

Larkey, P. D., Smith, R. A., & Kadane, J. B. (1989). It's okay to believe in the "hot hand". *Chance*, *2*(4), 22-30.

Lefgren, Lars, & Platt, B. (2010). "Coaches on the Hot Seat: Testing Models of Moral Hazard and Screening." Mimeo, Brigham Young University.

Lefgren, L., Platt, B., & Price, J. (2011). Sticking with what (barely) worked (Working Paper No. 17477). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w17477.pdf

Lefgren, L., Platt, B., & Price, J. (2014). Sticking with what (barely) worked: A test of outcome bias. *Management Science*. *61*(5*),*1121-1136.

Lipshitz, R. (1989). "Either a medal or a corporal": The effects of success and failure on the evaluation of decision making and decision makers. *Organizational Behavior and Human Decision Processes*, *44*(3), 380-395.

Marshall, Greg W., & John C. Mowen. (1993). An Experimental Investigation of the Outcome Bias in Salesperson Performance Evaluations. *Journal of Personal Selling & Sales Management*, *13*(3), 31-47.

Mackenzie, R., & Cushion, C. (2013). Performance analysis in football: A critical review and implications for future research. *Journal of Sports Sciences*, *31*(6), 639-676.

Martens, R., Vealey, R. S., & Burton, D. (1990). Development and validation of the competitive state anxiety inventory-2. In B. Owens (Ed.), *Competitive anxiety in sport* (Rev. ed., pp. 117-213). Champaign, IL: Human Kinetics Publishers.

Martin, J. J., & Gill, D. L. (1991). The relationships among competitive orientation, sport-confidence, self-efficacy, anxiety, and performance. *Journal of Sport & Exercise Psychology*, *13*, 149-159.

Mazzocco, P. J., Alicke, M. D., & Davis, T. L. (2004). On the robustness of outcome bias: No constraint by prior culpability. *Basic and Applied Social Psychology*, *26*(2-3), 131-146.

Mertins, L., Salbador, D., & Long, J. H. (2013). The outcome effect – A review and implications for future research. *Journal of Accounting Literature*, *31*(1), 2-30.

Miller, J. B., & Sanjurjo, A. (2014). A cold shower for the hot hand fallacy. IGIER working paper no. 518. SSRN Electronic Journal. Retrieved from http://ssrn.com/abstract=2450479.

Morris, B. (2015, February 2). A head coach botched the end of the Super Bowl, and it wasn't Pete Carroll. *Five Thirty Eight*. Retrieved from http://fivethirtyeight.com/features/a-head-coach-botched-the-end-of-the-super-bowl-and-it-wasnt-pete-carroll/

Moskowitz, T., & Wertheim, L. J. (2011). The myth of the hot hand. *Scorecasting: The hidden influences behind how sports are played and games are won* (pp. 215-229). New York, NY: Crown Archetype.

Nelson, L. J., & Groom, R. (2012). The analysis of athletic performance: some practical and philosophical considerations. *Sport, Education and Society*, *17*(5), 687-701.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, *2*(2), 175-220.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological review*, *84*(3), 231-259.

Ostrow, A. C. (2002). Anxiety. In S. R. Woods (Ed.), *Directory of psychological tests in the sport and exercise sciences* (2nd ed., pp. 47-97). Morgantown, WV: Fitness Information Technology.

Overman, S. J. (1999). 'Winning isn't everything. It's the only thing': The origin, attributions and influence of a famous football quote. *Football Studies*, *2*(2), 77-99.

Plessner, H., & Haar, T. (2006). Sports performance judgments from a social cognitive perspective. *Psychology of Sport and Exercise*, *7*(6), 555-575.

Puckett, A., (2015, February 2). Twitter reaction to Seahawks 4th-quarter play call. *ESPN*. Retrieved from http://espn.go.com/blog/nflnation/post/_/id/160678/twitter-reaction-to-seattle-seahawks-fourth-quarter-interception

Ratner, R. K., & Herbst, K. C. (2005). When good decisions have bad outcomes: The impact of affect on switching behavior. *Organizational Behavior and Human Decision Processes*, *96*(1), 23-37.

Reeves, J. L., Tenenbaum, G., & Lidor, R. (2007). Choking in front of the goal: The effects of self-consciousness training. *International Journal of Sport and Exercise Psychology*, *5*(3), 240-254.

Reyes, L. (2015, February 2). Stunned Seahawks still 'don't understand' final play call. *USA TODAY*. Retrieved from

http://www.usatoday.com/story/sports/nfl/2015/02/02/seahawks-super-bowl-final-play-call-interception-stunned-surprised/22729913/

Roese, N. J., & Vohs, K. D. (2012). Hindsight bias. *Perspectives on Psychological Science*, *7*(5), 411-426.

Rovell, D. (2016, May 12). Kobe inc. files to trademark Bryant nickname 'black mamba'. *ESPN*. Retrieved from http://www.espn.com/nba/story/_/id/15520468/kobe-bryant-company-files-trademark-black-mamba

Sacchi, S., & Cherubini, P. (2004). The effect of outcome information on doctors' evaluations of their own diagnostic decisions. *Medical Education*, *38*(10), 1028-1034.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social psychology*, *61*(2), 195-202.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *The journal of finance*, *53*(5), 1589-1622.

Solomonov, Y., Avugos, S., & Bar-Eli, M. (2015). Do clutch players win the game? Testing the validity of the clutch player's reputation in basketball. *Psychology of Sport and Exercise*, *16*, 130-138.

Ste-Marie, D. M., & Valiquette, S. M. (1996). Enduring memory-influenced biases in gymnastic judging. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1498-1502.

Taleb, N. N. (2007). The narrative fallacy. *The black swan: The impact of the highly improbable* (pp. 62-84). New York, NY: Random House.

Taylor, J. (1987). Predicting athletic performance with self-confidence and somatic and cognitive anxiety as a function of motor and physiological requirements in six sports. *Journal of Personality*, *55*(1), 139-153.

Tversky, A., & Gilovich, T. (1989). The "hot hand": Statistical reality or cognitive illusion?. *Chance*, *2*(4), 31-34.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, *76*(2), 105-110.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, *5*(2), 207-232.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453-458.

Vergin, R. (2000). Winning streaks in sports and the misperception of momentum. *Journal of Sport Behavior*, *23*(2), 181-197.

Whitley, B. E., & Kite, M. E. (2012). Survey research. In, *Principles of research in behavioral science* (3rd ed., pp. 436-479). New York, NY: Routledge.

Wilson, T. D. (2002). Knowing why. *Strangers to ourselves: Discovering the adaptive unconscious* (pp. 93-115). Cambridge, MA & London, UK: The Belknap Press of Harvard University Press.

Woodman, T., & Hardy, L. (2003). The relative impact of cognitive anxiety and self-confidence upon sport performance: A meta-analysis. *Journal of Sports Sciences*, *21*(6), 443-457.

Footnotes

[1]This participant was deemed to be an outlier because their difference score between round 1 and round 2 exceeded 2.5 standard deviations for male participants (*SD* = 3.03). In round 1 this participant made 5 free throws and in round 2 they made 13 shots. This corresponds to a difference score of 8 made baskets which was 2.64 standard deviations beyond the mean

[2]Using this abridged dataset, the other three hypotheses were tested through a series of t-tests. No statistically significant results were obtained and subsequent Cohen's *d* tests yielded no effect sizes of note.

[3]Analyses of variance (ANOVAs) were conducted to determine if skill level was differentially impacted by outcomes. As was expected given the small samples, no significant differences were revealed. In light of these sample size concerns, a means table (Table 1) is provided.

[4]For males only (i.e., the abridged dataset mentioned in footnote[2]), excluding the outlier (see footnote[1]), several different means on the basis of skill (i.e., highest level achieved) bear mention. Recreational players did not improve during round 2, but junior varsity players (M_{Diff} = 0.6) improved by over half a shot and varsity athletes improved by nearly a full shot (M_{Diff} = 0.82). As expected, varsity athletes made, on average, more shots (about 2 to 3 more shots) than junior varsity and recreational level athletes. Further, recreational and varsity level athletes rated their performance similarly following round 1, while junior varsity athletes provided lower performance ratings. Finally, varsity athletes tended to be more self-confident than junior varsity and recreational athletes.

Once again, because these figures represent mean scores only no casual inferences can be drawn.