THE GAMBLER'S FALLACY AND HOT OUTCOME: COGNITIVE BIASES OR ADAPTIVE THINKING FOR GOALKEEPERS' DECISIONS ON DIVE DIRECTION DURING PENALTY SHOOTOUTS

Abhishek Sarkar

A Thesis

Submitted to the Graduate College of Bowling Green State University in partial fulfillment of the requirements for the degree of

MASTER OF EDUCATION

May 2017

Committee:

David A. Tobar, Advisor
Adrian P. Turner
Matt Laurent
ABSTRACT

David A. Tobar, Advisor

In the face of uncertainty, human judgment and decision-making often tends to deviate from the realm of rationality. Gambler’s fallacy and its opposite, hot outcome, are two such departures from laws of probability involving random streak of events. However, the adaptive thinking approach to decision-making proposes that any irrational heuristic or illogical belief can be fully adaptive as long as it fulfills the requirements of the decision task. During penalty shootouts in association football, goalkeepers face a series of multiple penalty kicks which are independent draws from a random process, and they need to anticipate the likely kick directions with limited time, insufficient information and computational capacity. The objective of this current study was to observe the goalkeepers in real-world competitive settings and examine whether they use gambler’s fallacy and hot outcome as predictive strategies to decide on their dive directions during penalty shootouts following streaks of correct and incorrect predictions in the same direction. Another goal was to investigate from the adaptive thinking perspective whether such strategies lead to more correct predictions by the goalkeepers. Penalty shootout data were collected from the elite soccer tournaments over the course of last 25 years (1992 - 2016) and after applying appropriate exclusion criterions, 405 penalty kicks were considered for the final analysis. Binomial tests revealed that following progressively longer streaks of correct predictions in the same direction, goalkeepers became increasingly more likely to dive in the opposite direction for the subsequent kick than would be expected by chance, a behavior consistent with gambler’s fallacy. However, neither gambler’s fallacy nor hot outcome type patterns were observed when analyzing dives of goalkeepers following streaks of incorrect
predictions in a particular direction. On the other hand, results of the Fisher’s exact tests confirmed that both the fallacies failed to produce significantly more correct predictions from the goalkeepers and hence, they are mere fallacies. Goalkeepers’ belief in gambler’s fallacy highlights the biases associated with their real-world decision-making as this study does not support the adaptive use of such deceptive beliefs as predictive decision strategies.
To the Almighty, ‘the beautiful game’ and the memory of my elder brother, Jayanta Kumar

Bhowmick
I would like to offer my sincerest thanks and very profound gratitude to my advisor, Dr. David Tobar, my committee members, Dr. Adrian Turner and Dr. Matt Laurent, my school, School of Human Movement, Sport, & Leisure Studies at Bowling Green State University, my beloved parents, Mr. Arup Kumar Sarkar and Mrs. Kaberi Sarkar, and all my near and dear ones for their valuable guidance, inspiration, kind patience, constant encouragement and unconditional support throughout this endeavor. You all were there during my good and bad times and stood beside me when I needed you the most. Without you this journey would not have reached its destination successfully. I am forever indebted to you all. May God bless you!

I would also like to acknowledge the help by Dr. Ignacio Palacios-Huerta of London School of Economics for answering my questions and explaining his research every time patiently.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>LITERATURE REVIEW AND RESEARCH OBJECTIVES</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Gambler’s Fallacy and Hot Outcome</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Adaptive Thinking Approach to Decision-Making and Adaptiveness of Incorrect Beliefs</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Goalkeepers’ Strategies for Predicting Penalty Kick Directions</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Cognitive Biases in Goalkeepers’ Actions</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Summary of Literature Review</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Research Objectives</td>
<td>20</td>
</tr>
<tr>
<td>III</td>
<td>METHODS</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Analysis Methods</td>
<td>24</td>
</tr>
<tr>
<td>IV</td>
<td>ANALYSIS AND RESULTS</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Descriptive Analysis</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Gambler’s Fallacy and Hot Outcome</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Biases or Behaviors</td>
<td>34</td>
</tr>
<tr>
<td>V</td>
<td>DISCUSSION</td>
<td>38</td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
<td>43</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Notions of human rationality and decision-making approaches</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>Percentage of goalkeepers’ dives in the opposite or same direction in the next kick after streaks (of correct predictions in the same direction) of varying lengths</td>
<td>31</td>
</tr>
<tr>
<td>3</td>
<td>Percentage of goalkeepers’ dives in the opposite or same direction in the next kick after streaks (of incorrect predictions in the same direction) of varying lengths</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td>Correct prediction probabilities for gambler’s fallacy and hot outcome dives after each streak (of correct predictions in the same direction) length</td>
<td>35</td>
</tr>
<tr>
<td>5</td>
<td>Correct prediction probabilities for gambler’s fallacy and hot outcome dives after each streak (of incorrect predictions in the same direction) length</td>
<td>36</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Number and Percentage Share of Penalty Shootouts in the Dataset</td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>Examples of Gambler’s Fallacy and Hot Outcome Dives Following One Correct or Incorrect Predictions with Any Particular Direction (Left or Right) in the Data</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>Examples of Gambler’s Fallacy and Hot Outcome Dives Following Two Consecutive Correct or Incorrect Predictions in the Same Direction (Left or Right) in the Data</td>
<td>26</td>
</tr>
<tr>
<td>4</td>
<td>Examples of Gambler’s Fallacy and Hot Outcome Dives Following Three Consecutive Correct or Incorrect Predictions in the Same Direction (Left or Right) in the Data</td>
<td>26</td>
</tr>
<tr>
<td>5</td>
<td>Frequency of Goalkeepers’ Dives in the Opposite or Same Direction in the Next Kick after Streaks (of Correct Predictions in the Same Direction) of Varying Length</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>Frequency of Goalkeepers’ Dives in the Opposite or Same Direction in the Next Kick after Streaks (of Incorrect Predictions in the Same Direction) of Varying Length</td>
<td>32</td>
</tr>
</tbody>
</table>
CHAPTER I. INTRODUCTION

The main factor in a penalty shootout is luck again. You need to stay calm and focused but the biggest thing you need is luck.

-Peter Shilton (Former England Goalkeeper), *MHF*, 2006

In spite of its ill repute as an ‘unfair’ affair in ‘the beautiful game,’ the penalty shootout in association football (soccer) offers a unique platform to explore and analyze real-world sequential performance and decision-making under uncertainty. Designed for the sake of determining a winner after failing to identify such at the end of the extra time in a knock out match, the penalty shootout (or kicks from the penalty mark) is based on a sequence of multiple penalty kicks until one team has a greater number of goals after five kicks per team or one team has a one goal advantage over the other after the same number of kicks in ‘sudden death’ (FIFA, 2016). A fascinating study by Palacios-Huerta (2002) has revealed that the sequences of the kick direction (left or right) generated by the kickers during a penalty shootout truly emerge from an independent and identically distributed (i.i.d.) random process. That means, kickers’ choices are random and serially independent and hence, goalkeepers’ success probabilities are statistically indistinguishable across both dive directions (left or right). Such behavior on the part of human beings is exceptionally rare as typically we are sub-standard when it comes to act or perform in random manner (Wagenaar, 1972; Treisman & Faulkner, 1990). With no evidence that the goalkeepers are more likely to anticipate the kick directions accurately than would be expected by chance, their behavior suggests that they “endogenously” select dive directions (Misirlisoy & Haggard, 2014). Hence, it is reasonable enough to compare the event of prediction of kick directions by the goalkeepers during a penalty shootout with that of speculating outcomes in a
series of independent binary random events, such as coin tossing and gambling.

Humans are intrinsically poor at understanding or interpreting chance (Bocskocsky, Ezekowitz, & Stein, 2014). We generally seek for pattern and regularity amidst randomness, and while doing so, we often rely on deceptive and illogical intuitive ideas or heuristics which lead to systematic deviation from rationality (the laws of chance). Tversky and Kahneman (1971), in their pioneering work, attributed such cognitive or perceptual biases to local representativeness, an intuition based on the law of small numbers, whereby people wrongly perceive that “the essential characteristics of a chance process to be represented not only globally in the entire sequence, but also locally, in each of its parts” (Gilovich, Vallone & Tversky, 1985, p. 296).

Gambler’s fallacy and hot outcome are examples of two erroneous beliefs based on the assumption of local representativeness. Gambler’s fallacy is defined as a (mistaken) belief in “negative autocorrelation of a non-autocorrelated random sequence of outcomes” (Sundali & Croson, 2006, p. 2). That is, a streak of outcomes generated by a sequence of independent random draws become increasingly likely to break as the streak becomes progressively longer. For example, a person bets on a lucky side every time a fair coin is flipped (or different coins with identical outcome distribution are flipped repeatedly) and wins on every attempt. Then, if that person believes in gambler’s fallacy, he or she will predict that the lucky side will be less likely than chance to come up on the next toss and thus call for the opposite side. Although the events are independent and have no causal relation. Hot outcome is the opposite of gambler’s fallacy (Sundali & Croson, 2006). It is a (mistaken) credence in “positive autocorrelation of a non-autocorrelated random sequence” (Sundali & Croson, 2006, p. 2). That is, a streak of outcomes generated by a sequence of independent random draws become increasingly likely to continue as the streak becomes progressively longer and that outcome turns ‘hot.’ For example,
following a series of successful predictions while betting on a lucky side before every time an unbiased coin is tossed (or multiple coins with identical outcome distribution are tossed repeatedly), a person believing in hot outcome will perceive that the lucky side will be more likely than chance to come up on the next toss and thus opt for the same side even though the events are uncorrelated and serially independent. It is to be noted that there is a conceptual difference between hot outcome and the famous ‘hot hand’ belief. In hot hand belief, it is perceived that the person is hot whereas in hot outcome, the outcome of the random process is judged to be hot.

It has been a matter of longstanding discourse whether an irrational belief can be modeled as adaptive behavior or not. Rationality of beliefs was the cornerstone of heuristics and biases approach of decision-making introduced by Kahneman and Tversky (Gilovich, Griffin, & Kahneman, 2002; Kahneman, Slovic, & Tversky, 1982). Gilovich (2002) supported this view while commenting that “I don’t see what is adaptive about the use of an outright fallacy”. However, with the emergence of the adaptive thinking (Anderson, 1990; Gigerenzer & Todd, 1999; Gigerenzer, 2000) decision-making approach, the focus shifted to the goal attainment credibility (i.e., whether the behavior can improve the outcome of the decision-making or not). Burns (2004) argued that “the fact that the behavior may be supported by a false belief is irrelevant to whether the behavior is adaptive or not, though the false belief may actually be beneficial to the extent to which it helps to maintain the adaptive behavior” (p. 296). In fact, researchers found concrete evidence that the use of irrational heuristics can lead to better judgement and thus surely be adaptive.

The objective of this research is to detect any apparent pattern of serial dependence (in the form of gambler’s fallacy and hot outcome) in goalkeepers’ real-world choices when facing
sequences of independent random events and also to analyze whether belief in such statistical illusions is mere bias or can be adaptive. This study, using empirical data from elite soccer tournaments over the last 25 years, will investigate whether professional goalkeepers display behavioral patterns consistent with the gambler’s fallacy and hot outcome while judging kick directions following repeated correct or incorrect predictions in the same direction during penalty shootouts. Another goal is to examine whether or not the belief in such irrational sequential patterns leads to more correct predictions on the part of the goalkeepers.

Penalty shootouts are of great importance during the elimination stages of the major tournaments, and they are highly frequent too. For example, in the last 10 years, 4 UEFA Champions League finals have been decided in penalty shootouts. The presence of strict rules (e.g., goalkeepers can only move sideways during the kick - they are not allowed to come forward from the goal line.) imposed by the rules of the game and the unavoidable physical (O’Callaghan, 2014), social (Lopes, Araujo, Peres, Davids & Barreiros, 2008) and psychological (Apesteguia & Palacios-Huerta, 2010; Wood, Jordet & Wilson, 2015) constraints, make a penalty kick an arduous task for the goalkeepers. They must anticipate the likely kick direction and dive promptly so that they have a chance of reaching the ball to make a save. Goalkeepers often do not have sufficient information on which to base their decision for choosing the dive direction, thus making it an event full of uncertainty and chance. Also, since the penalty kicks have two possible outcomes only: goal or no goal, the cost of committing an error is huge and such an error becomes more statistically significant during the penalty shootouts (McGarry & Franks, 2000). Moreover, saving a penalty kick or a few kicks during a penalty shootout serves as a significant psychological boost for the remaining kickers of the goalkeeper’s team as it eases the pressure on them. Therefore, the first and foremost task for a goalkeeper during the penalties
is to correctly judge the direction of the kick. This study will observe and analyze whether
goalkeepers do believe in the streaks (i.e., whether they go against the streaks- gambler’s fallacy
and/or they follow the streaks- hot outcome) in order to find meaningful patterns in random
occurrence of kick directions during penalty shootouts and also whether or not going against the
streaks and following the streaks can indeed be beneficial for them. The implication of this study
is vital as it intends to identify any irrational sequential patterns or biases emerging from the
goalkeepers’ behavior in real-world competition (i.e., how goalkeepers actually make decisions
in reality). Moreover, if these biases can be found to yield significantly better prediction
outcomes, this study would be the first to provide goalkeepers with some cognitive strategies, or
fast and frugal heuristics, based on the sequential nature of the penalty shootouts to more
successfully decide on their dive direction.
CHAPTER II. LITERATURE REVIEW AND RESEARCH OBJECTIVES

Gambler’s Fallacy and Hot Outcome

The gambler’s fallacy and hot outcome illustrate human misconceptions of non-autocorrelated random sequences and has been part of extensive research in the fields of psychology, neuroscience, economics, gambling games, etc. Gambler’s fallacy, or the negative recency effect, arises when “evaluating sequences generated by a random mechanism, people believe that streaks of events will be shorter than they are in true Bernoulli binomial sequences” (Oskarsson, Van Boven, McClelland, & Hastie, 2009, p. 262). Gambler’s fallacy is thought to be caused by representativeness bias (Tversky & Kahneman, 1971; Kahneman & Tversky, 1972). According to this belief, people perceive that “short random sequences should reflect (be representative of) the underlying probability used to generate them,” (Croson & Sundali, 2005, p. 197) and they see chance as “a self-correcting process in which a deviation in one direction induces a deviation in the opposite direction to restore the equilibrium” (Tversky & Kahneman, 1974, p. 1125). For example, in a coin flipping scenario, a person believing in gambler’s fallacy will call for a tail following two consecutive heads because of his misconception about the representativeness of the underlying distribution (i.e., the sequence head-head-tail is more representative than head-head-head as it is wrongly perceived that the distribution of head and tail is always 50:50). Mullainathan (2002) in his paper explained this phenomenon more explicitly by asserting that “people often demonstrate such a predilection for expecting mean reversion, for expecting sequences to right themselves out” (p. 19). That is, after two heads in a row, a tail is more justifiable as it will make the sequence more representative of its parent population. Rabin (2002) explained this phenomenon by highlighting the notion that people generally overestimate the extent to which the small samples represent the parent populations.
and hence, a person who believes in the gambler's fallacy would think that early streak of one outcome increases the odds of other outcomes at the subsequent draws.

The French mathematician and astronomer Pierre-Simon Laplace (1825) provided the first insight into gambler’s fallacy in his essay, “Illusions in the Estimation of Probabilities,” drawing examples from gambling, lotteries, etc. Hans Reichenbach (1934/1949) also asserted the existence of the similar kind of patterns in people’s behavior during events like the coin toss and roulette spins where they expect the sequences of outcomes to alter too frequently than they do and shift their bets accordingly. Since then, several behavioral studies have investigated the existence of gambler’s fallacy or negative recency-type pattern in prediction, and most of these studies provide evidence in support of its presence. This effect was first observed during experiments on ‘probability matching’ in laboratory settings (under controlled conditions) where participants were instructed to predict the color of the light that would illuminate next (Sundali & Croson, 2006). Remarkably, following a streak of one particular color of light, the participants demonstrated a significant tendency to opt for the opposite color, a behavior that is referred to as negative recency in the literature (e.g., Estes, 1964; Lee, 1971). Several other researchers have also found evidence of gambler’s fallacy type beliefs during laboratory experiments. Ayton and Fischer (2004) analyzed bets placed by the participants predicting the color outcome on a roulette wheel and observed negative recency in subjects’ expectations following streaks of outcomes of length 1-5 although the sequence of outcomes are equally likely and independent. Gal and Baron (1996) in their experiment asked the participants to generate or evaluate different strategies to bet on the most likely outcome (red or green, differing in probability) on repeated trials of a die rolling experiment and found evidence for the belief in gambler’s fallacy in their behavior. Boynton (2003) performed an experiment in which he asked the subjects to predict the
next outcome after looking at a random binary sequence. He divided the subjects into 3 groups – one of them was given no information about the sequence, one of them was told that the sequence was generated from a random event like a coin toss, and the final group was told that another student has prepared the sequence. The group that was told it was a random event showed significant gambler’s fallacy-type behavior that was only minimally evident in the other groups. Gold (1997) also detected similar kind of pattern during series of coin tosses where the participants committed gambler’s fallacy. However, the most interesting finding of this study was that people demonstrated less gambler’s fallacy-type behavior when multiple dissimilar coins (e.g., a cent and a quarter) were tossed than in the case of similar type multiple coins.

Other studies have provided empirical evidence of the existence of gambler’s fallacy. Terrell and Farmer (1996) and Terrell (1998) studied betting patterns in dog racing and found strong evidence of gambler’s fallacy as the bettors “underestimate the probability of a repeat win by the dog in the post position of the last winner” (p. 162). If a dog in post-position 1 won the previous race, then the bettors would be less likely to bet on the dog (a different one) in the same post-position in the next race. After examining bets on 12,316 horse races, Metzger (1985) asserted the existence of gambler’s fallacy-type bias among the bettors as they were significantly more reluctant to bet on the favorite horse (different horse) given that the favorites have won repeatedly in the recent past. Many studies have confirmed the belief of gambler’s fallacy in lottery play prediction. Clotfelter and Cook (1993) found that the number of bets on the winning number fall drastically “after they were drawn to approximately 60% of their previous popularity and then gradually returned to their original ambient popularity after 3 months” (Oskarsson et al., 2009, p. 266), although the outcomes are serially independent and uncorrelated across days. Terrell (1994) replicated Clotfelter and Cook’s (1993) study using data from pari-mutuel lottery
where they used player winnings as an index of under and over betting. In pari-mutuel lottery, all the bets of similar type are pooled together and payoff odds are calculated by dividing the pool equally among all winning bets. Their study also confirmed the same kind of expectation as approximately 25% fewer bettors bet on a winning number from the previous week than on a number which had not won in the last 8 weeks. Gambler’s fallacy-type behavior was also confirmed in birth-sex prediction studies. McClelland and Hackenberg (1978) asked people to predict the sex of the next child after informing them about the sex of the existing children in a family (e.g., 2 boys and 3 girls). It was found out that most of the subjects “predicted that a family’s sex composition would balance out such that the less numerous sex would be the more likely sex of the next child” although such birth dependence is trivial (Oskarsson et al., 2009, p. 267). Croson and Sundali (2005) tested for the existence of such statistical illusions in casino gambling using individual field data and found out that following 5 or more repeated occurrences of a particular outcome, the individuals become progressively more likely to bet against that outcome. A study by Xu and Harvey (2014), with data from online gambling, revealed that although there is no existence of gambler’s fallacy streaks (i.e., following losses, gamblers are more likely win), the gamblers opted for safer odds after winning consecutive bets and selected riskier ones following losses on successive bets (i.e., they behaved in accordance with the gambler’s fallacy). In professional sport settings, Walker and Wooders (2001) found evidence of negative autocorrelation in the choices of first serves performed by the professional tennis players (i.e., whether they went to the forehand or backhand of the opponent) even though the optimal strategy for them is to act randomly so that their actions do not become too predictable to the opponent. Kovash and Levitt (2009) studied more than three million pitches in baseball and 125,000 play choices from NFL and in both sports found clear negative serial correlation in
The opposite of gambler’s fallacy is the hot outcome which can be viewed as one’s wrong perception that a random sequence will exhibit positive recency, or “more generally, an incorrect belief in a more positive autocorrelation than is present” (Sundali & Croson, 2006, p. 2). Stemming from the same source like gambler’s fallacy, under this assumption, people wrongly overestimate the degree of representativeness of small samples to their population. Hence, for example individuals who believe in hot outcome will believe that after three heads, another head is more likely to appear than a tail because the coin is biased toward the ‘hot’ outcome head.

Edwards (1961), Lindman and Edwards (1961), and Feldman (1959) found evidence favoring hot outcome during probability matching experiments in laboratory settings. Following a streak of one outcome, the subjects were significantly more likely to guess the same outcome in the subsequent turn. Ayton and Fischer (2004) observed that although people demonstrated negative recency while predicting the color outcomes on a simple roulette wheel, but they expected positive recency for the sequence of successes and failures of their predictions. Karen and Lewis (1994) also asserted the presence of a positive recency effect in gambling type tasks in the lab. They found that when asked to predict the minimum number of observations required to detect a particular hot outcome (number) on a roulette wheel, they typically underestimate the observations required. “Thus after seeing even a small streak of red numbers, gamblers might believe the wheel is biased and expect more red numbers” (Sundali and Croson, 2006, p. 3). However, their construct is not exactly identical to the hot outcome, and empirical studies supporting the belief in hot outcome are extremely rare in the existing literature. Sundali and Croson (2006) provided evidence for the existence of the positive recency effect by analyzing 18
hours of casino data where a significant number of bettors demonstrated a tendency to bet on an outcome that has been observed recently. Some bettors believed in hot outcome and other bettors displayed gambler’s fallacy type behavioral pattern. Their findings suggest that there are significant individual differences in the perception of randomness.

**Adaptive Thinking Approach to Decision-Making and Adaptiveness of Incorrect Beliefs**

Human decision-making models rests firmly on the notion of human rationality. It typically comes in two forms: unbounded rationality and bounded rationality. Unbounded rationality “encompasses decision-making strategies that have little or no regard for the constraints of time, knowledge and computational capacities that real humans face” (Gigerenzer & Todd, 1999, p. 7). From this perspective, decision-making can be viewed as a fully rational process or seeking an optimal choice and it is typically modeled by probability theory. The concept of bounded rationality or limited rationality was originally introduced by economist Herbart Simon (1955, 1957) on the grounds that human decision tasks are immensely complex to optimize with limited time and resources. Some tasks may not also have any perfect “calculable optimal solution” (Bennis & Pachur, 2006). Simon concluded that while making decisions under such scenarios, our rationality is limited by the complexity of the task, insufficient time and computational limitations of our minds. He identified a heuristic – ‘Satisficing’ – a simple rule of thumb that involves selecting the first choice that satisfies “one’s minimum criteria” (Bennis & Pachur, 2006, p. 614).

Resting on the assumption of unbounded rationality, the objective of the normative or prescriptive approach to decision-making is to model the best ‘rational’ solution with perfect accuracy while avoiding the real-world constraints. It is mostly concerned with what people
should do rather than focusing on what real people actually do while making decisions. Although the concept was initially coined by Simon, the study of heuristics as human decision-making tools was developed by psychologists Tversky and Kahneman with the intention of “identifying the beliefs people hold, often by making inferences from apparently incorrect behavior” (Burns, 2004, p. 326). Using the heuristics and biases approach, they identified several heuristics (or practical problem-solving methods) using their own perspective of bounded rationality (e.g. availability, anchoring and adjustment, representativeness) and deduced the general biases associated with them (e.g., deviation from laws of probability). They commented that although these heuristics may be “serviceable”, they are not accurate. These mental short cuts can lead to systematic deviation from the realm of rationality and wrong conclusions (Gilovich, Griffin & Kahneman, 2002). They termed such beliefs as mere cognitive biases or fallacies. Gerd Gigerenzer and others have strongly criticized the heuristics and biases approach for emphasizing the fact that how heuristics lead to errors. Gigerenzer and Todd (1999) suggested that the objective should be to attain the goal of satisfying the requirement of the decision tasks and improve decision-making by staying within the bounds of real world constraints. Adaptive thinking approaches to decision-making focuses on using simple heuristics to achieve satisfactory solutions while not being precise entirely. Gigerenzer and Todd (1999) commented that people use an ‘adaptive toolbox’ to base their reasoning for their decisions, which can be defined as “the collection of specialized cognitive mechanisms that evolution has built into the human mind for specific domains of inference and reasoning, including fast and frugal heuristics” (p. 30).
Figure 1. Notions of human rationality and decision-making approaches.

Both the heuristics and biases and the adaptive thinking approaches rely on the notion of bounded rationality and encourage the use of heuristics over the normative optimizing rational choice models to facilitate decision-making process, but there lies some fundamental differences from the angle of rationality and adaptiveness. The objective of the heuristics and biases approach is to find the analytically correct choices and in this sense, it is very close to the normative approach. On the other hand, the adaptive thinking approach to judgement and decision-making rests firmly on the ‘principles of adaptivity’ and tends to ignore the fact of rationality or irrationality of a belief, but rather focuses on whether it can improve human decision-making or not. Thus, as long as a behavior supported by a logically irrational belief yields fruitful results, it may be considered adaptive whereas such irrational heuristics are not adaptive under the heuristics and biases approach. Anderson (1990) supported this view when he commented that ‘‘there is no reason why normatively irrational heuristics cannot be adaptive’’
The study of the ‘hot hand’ phenomenon in the game of basketball illustrates the significant distinction between rationality of a belief and adaptiveness of a behavior. Gilovich and colleagues (1985) analyzed professional basketball players’ successive shots and found no evidence of any dependence. They concluded that hot hand is a mere reasoning misconception. The basketball players wrongly perceived that ‘success breeds success’ when they gave the ball to a player who has enjoyed a streak of successful shots. But, “this conclusion is based on the implicit assumption that showing a belief (dependence) to be a fallacy is equivalent to showing a behavior (use streaks as an allocation cue) is also invalid” (Burns, 2003, p. 193). However, such assumption cannot be assumed. From the perspective of adaptiveness, the more relevant question to ask in basketball is whether belief in the hot hand can lead to a higher team score, rather than focusing on the notion of incorrectness of the belief (i.e., independence of successive shots).

“Thus there is a hot hand belief and a hot hand behavior. The distinction between the hot hand behavior and belief is critical to evaluating the adaptivity of positive recency in basketball as a heuristic” (Burns, 2003, p. 193). Burns (2001, 2004) provided empirical evidence in favor of the hot hand behavior. He used simulations and markov models to show that streaks are “predictive of a player’s shooting percentage” and thus serve as a valid cue for identifying the best shooters. He demonstrated that passing the ball to a ‘hot’ player while ignoring the players’ base rates attains the goal of scoring more and thus increasing the probability of winning. Hence, although such belief is incorrect, it can be a fast and frugal heuristic for the basketball players. Arkes (2010) also provided evidence on the adaptiveness of hot hand in basketball when he examined all free throws during the 2005-06 NBA season with a multivariate framework with individual fixed effects and found evidence for the hot hand in that making the first free throw is associated
with a significantly higher probability of making the second free throw.

Other studies have also found that the hot hand can indeed be adaptive. Raab, Gula, and Gigerenzer (2012) found that hot hand streaks do exist in the game of volleyball, and players and coaches can detect it and use it as an adaptive behavior. They found that believing in the hot hand led to more hits for a team. A study on professional golfers (Livingston, 2012) has also revealed that there is an impact of ‘good’ streaks on the performance of the players, but its influence can vary with the players’ level of experience.

In a study by Xu and Harvey (2014), hot hand streaks were found to exist in online gambling. Gamblers were more likely to win following a streak of wins suggesting that they should keep betting. However, surprisingly they found that it was the belief in gambler’s fallacy that led to the creation of hot streaks. Following a win, they strategically opted for less risky odds whereas after a loss, they chose less safe odds, a behavior which is consistent with the gambler’s fallacy. The findings of this study is particularly significant as it is one of the few studies that shows that belief in gambler’s fallacy can also be adaptive. Hence, previous research supports the validity of the adaptive thinking approach.

**Goalkeepers’ Strategies for Predicting Penalty Kick Directions**

During a penalty shootout, a goalkeeper faces multiple kicks in quick succession unlike match-play penalty kicks which are typically separated by large time lags. Penalty kicks illustrate the complexity and ambiguity of real world sport decision tasks. The decision-making process has always been an area of high interest among researchers from diversified fields, such as sports science, psychology, neuroscience, economics and even mathematics. Both the actors (kickers and goalkeepers) directly involved in the task must base their judgement on insufficient
information and few available strategies to make a decision within a very limited amount of time. However, this study will only focus on the goalkeepers and existing research on their strategies to anticipate the kick directions.

A penalty kick is shot from a distance of 36 feet from the goal line, and on average, the ball travels with a speed of 70 mph and takes less than half a second to cross the goal line (O’Callaghan, 2014). Also, the large size of the goal (height - 8 feet and width - 24 feet) contributes to the difficulty of the goalkeeper’s task leaving insufficient time for the goalkeeper to observe the direction of the ball and move the body in that direction. Hence, the goalkeepers must anticipate the likely kick direction before the ball is actually kicked (Chiappori, Levitt, & Groseclose, 2002; Palacios-Huerta, 2003). This puts the odds overwhelmingly in a kicker’s favor. The goalkeepers have very few strategies available in their inventory to base such decisions and exploration of such strategies has been a topic of extensive research. Studies on visual search behavior and anticipatory skills have revealed that looking for “advanced postural cues,” such as hips and non-kicking foot of the kickers, can help in predicting the direction of the kicks (Morris & Burwitz, 1989; Savelsbergh, Williams, Van der Kamp, & Ward, 2002; Savelsbergh, Van der Kamp, Williams & Ward, 2005). The researchers found that expert goalkeepers are more efficient in observing these cues and anticipating the direction of the kicks accurately more often than the novices. Other studies have focused on identifying goalkeeping strategies that can influence the kick direction. Masters, Van der Kamp, and Jackson (2007) and Noel, Kamp, Weigelt, and Memmert (2015) found that imperceptibly off-centered goalkeepers (e.g., 6-9 cm to left or right) can significantly influence the direction of the kicks. The results revealed that although the kickers typically fail to identify such displacement, they were more prone to shoot on the larger, more open side of the goal. In another study, Wood and Wilson
(2010) showed that a moving goalkeeper (e.g., waving arms) creates a notable distraction for the kickers and results in more saves. The kickers find it difficult to shift their attention away from the goalkeeper’s movements, and it negatively affects their shooting strategies and accuracy. Although there exist a substantial amount of literature on goalkeepers’ decision-making strategies to predict the penalty kick directions, almost all of the research has been carried out in a laboratory setting under controlled conditions and was focused on addressing ‘ideal’ behavior on the part of the goalkeepers (i.e., what goalkeepers should do) while ignoring the constraints that the goalkeepers face in real-world such as time, complexity of the task and computational limitations. Hence, this research lacks ecological validity on how goalkeepers actually make decisions in real-world settings and the effectiveness of the cognitive strategies they use.

Cognitive Biases in Goalkeepers’ Actions

This chapter discusses the real-world actions and biases of the goalkeepers during penalty kicks in soccer, interpreted within the framework of cognitive psychology and behavioral economics. In repeated one-shot two-person zero-sum strategic interaction games like penalty kicks, the most rational strategy is to act in an unpredictable manner to out-guess the opponent. Chiappori et al. (2002) and Palacios-Huerta (2003) studied the behaviors of goalkeepers in match-play penalty kicks and found that they play in accordance with the mixed-strategy Nash equilibrium (MSNE) predictions. In MSNE, a player adopts mixed strategy by assigning certain probabilities to each available pure strategy and then randomize between them. Chiappori et al. described the penalty kick as a simultaneous 3 * 3 game in which each player has three possible strategies (left, center, or right) and showed that their data support several predictions that should hold if players are playing MSNE. Palacios-Huerta analyzed the penalty kicks from the
perspective of 2 * 2 games (left or right only) and found that the goalkeepers’ choices are random and serially independent. Later, Palacios-Huerta (2002) confirmed that the same tendency also appears during the penalty shootouts. These studies could not detect any biases on the part of the goalkeepers.

However, more recent studies have found evidence of the presence of cognitive biases in goalkeepers’ behavior. Bar-Eli, Azar, Ritov, Keidar-Levin and Schein (2007) found that elite goalkeepers always prefer to jump left or right rather than staying in the center, thus suffering from a bias toward action. They attributed such behavior of the goalkeepers to norm theory. Since the norm is to jump, allowing a goal by being inactive (standing in the center) will feel worse than being active (jumping). Consequently, goalkeepers often choose a sub-optimal strategy rather than the optimal one, which is to stay in the center. However, such a result does not directly contradict the findings of Chiappori et al. (2002) and Palacious-Huerta (2003) as their analysis is based on the choices between center and the sides where as Chiappori et al. and Palacious-Huerta examined the choices between left and right. Roskes, Sligte, Shalvi, and De Dreu (2011) examined the FIFA World Cup penalty shootouts and discovered a surprising pattern in the goalkeepers’ behavior. The goalkeepers were significantly more likely to dive to the right than to the left when their team was losing, which they interpreted as being more approach motivated. “Approach motivation, a focus on achieving positive outcomes, is related to relative left-hemispheric brain activation, which translates to a variety of right-oriented behavioral biases” (Roskes et al., 2011, p. 1403). Since penalty kickers shot towards both directions equally often, this was a sub-optimal strategy (directional bias) on the part of the goalkeepers and they conceded more goals. However, their findings were also put into question by other researchers. Price and Wolfers (2014), following their analyses on a new set of data,
contradicted the findings of Roskes et al. (2011). They commented that there is no evidence of a significant presence of such directional biases in their data. Price and Wolfers showed that the behavioral pattern that emerged in Roskes et al.’s (2011) study was not a dysfunctional bias, but rather a perfectly rational strategy as kickers are more likely to kick on the right in these situations. Misirlisoy and Haggard (2014) were the first to investigate the existence of any sequential behavioral pattern that the goalkeepers might display during the penalty shoot outs and found that goalkeepers’ behave in accordance with gambler’s fallacy. Following repeated kicks in the same direction, goalkeepers become increasingly more likely to dive in the opposite direction in the subsequent kick. However, recently Braun and Schmidt (2015) critically analyzed the data used by Misirlisoy and Haggard and found that they used inappropriate methodology to determine the presence of gambler’s fallacy. A re-analysis of the same data along with new empirical and experimental data found that although goalkeepers do have a tendency to dive in the opposite direction of the last kick, there was no significant support of the gambler’s fallacy. These studies have analyzed the goalkeepers’ real-world behavior and explored the biases in their judgement. However, they were restricted to the to the identification of the biases or the errors that goalkeepers commit while making a decision; they did not intend to determine whether or not such biases can be effective behaviors also from the perspective of adaptive thinking approach. In sum, the results of these investigations are equivocal.

**Summary of Literature Review**

Previous research within the realm of adaptive thinking have strongly supported the notion that under certain circumstances, incorrect or irrational beliefs can indeed yield better results and thus be adaptive. Gambler’s fallacy and hot outcome are surely statistical illusions
(deviation from the laws of probability) regarding streaks emerging from series of purely independent and uncorrelated binary random trials. However, it is a distinct question that whether they can be used adaptively in prediction or not. Although existing literature has highlighted several strategies or biases to improve the anticipation skills and decision-making behavior of the goalkeepers, they still remain ‘lotteries’ to the goalkeepers, an event of pure luck and uncertainty. There is also not enough research evidence on whether goalkeepers display patterns similar to some of the erroneous beliefs while making decisions on dive directions during a penalty shootout and whether those beliefs are adaptive (i.e., can they generate better predictions?). Examples of studies looking into how goalkeepers actually make decisions in real-world are also extremely rare. Moreover, the adaptive thinking approach is yet to be tested in the context of soccer penalty shootouts.

**Research Objectives**

Penalty shootouts in soccer, due to its sequential nature and characteristics, provide a rare opportunity to study human being’s propensity towards gambler’s fallacy and hot outcome in real-world competitive setting. Henceforth an important question is whether these fallacious beliefs are mere judgement biases (i.e., incorrect choices) or can indeed produce better prediction outcomes.

The objectives of this research are twofold:

1. Do elite goalkeepers exhibit patterns in accordance with the gambler’s fallacy or hot outcome while anticipating kick directions following repeated correct or incorrect predictions in the same direction during penalty shootouts?

This study will investigate the existence of the gambler’s fallacy and hot outcome
effect in goalkeepers by examining the impact of previous prediction outcomes on their current prediction. Thus, in similarity to the results of the lottery (e.g., Clotfelter and Cook, 1993) and dog and horse racing (e.g. Metzger, 1985; Terrell, 1998) studies, a negative relationship between the previously correct prediction outcomes (or a positive relationship between the previously incorrect prediction outcomes) and the current prediction will support the existence of the gambler’s fallacy effect in goalkeepers (Croson & Sundali, 2005). On the other hand, a positive relationship between the previously correct prediction outcomes (or a negative relationship between the previously incorrect prediction outcomes) and the current prediction will provide evidence for the hot outcome.

2. Does belief in the gambler’s fallacy or hot outcome lead to more correct predictions on the part of the goalkeepers?

The studies by Misirlisoy and Haggard (2014) and Braun and Schmidt (2015) have already looked at the goalkeepers’ proneness to gambler’s fallacy. However, they did not examine it from the perspective of this study. Their construct was focused toward considering all the previous repeated kicks in the same directions irrespective of goalkeepers’ successes or failures (correct or incorrect predictions), whereas this study will only pick those repeated kicks: when the goalkeeper had a series of correct predictions in the same direction and when he had repeated incorrect predictions with the same direction. A study from this perspective can add to the knowledge base beyond the findings of Misirlisoy and Haggard (2014) or Braun and Schmidt (2015). Also, this study will be the first of its kind to examine the existence of the hot outcome effect among goalkeepers during the penalty shootout and investigate the adaptiveness (or mal-adaptiveness) of both the fallacies from the perspective of adaptive thinking approach. If existence of both or either one of the fallacies can be supported, that will also add to the existing
literature beyond the findings of Palacios-Huerta (2002) who showed that in general
goalkeepers’ choices are uncorrelated and serially independent during the penalty shootouts. It is
to be noted that the scope of this study is limited to prediction of kick directions only. Saving a
penalty kick depends on many factors such as timing of the movement and jump (Miller, 1998),
ball’s speed and trajectory, etc., and not only on the successful prediction of the kick direction.

The findings and implications of this study can vary widely depending on the results
obtained. If the results support that goalkeepers subscribe to the gambler’s fallacy and/or hot
outcome (consciously or sub-consciously) and those beliefs do not produce significantly more
correct predictions, then they would be considered as mere illusions only and goalkeepers should
not make decisions based on them. But, if it produces a significantly higher number of correct
predictions, then in accordance with the adaptive thinking approach, it should be considered as
an adaptive cognitive behavioral strategy which goalkeepers can use ‘adaptively’. On the other
hand, if it is observed that goalkeepers do not exhibit patterns in accordance with gambler’s
fallacy and/or hot outcome and those fallacies do not yield more accurate prediction outcomes
consistently, then it is to be believed that they are not prone to such sequential biases. However,
if such predictions produce better results consistently, then gambler’s fallacy or hot outcome can
be suggested as a fast and frugal heuristic for the goalkeepers to predict the kick directions
during the penalty shootouts.
CHAPTER III. METHOD

Data

The investigator filed a review determination application for the Human Subject Review Board (HSRB) of Bowling Green State University, and HSRB determined that this project did not meet the definition of human subject research and thus was exempt from full board review. The data for this research was collected from open source online videos (e.g., YouTube) and soccer statistics websites (e.g., Goal.com). These sites typically possess all the records, including the penalty shootouts, from the elite professional soccer tournaments. The researcher examined all the available penalty shootouts from the men’s FIFA World Cup, UEFA European Championship and UEFA Champions League tournaments over the course of last 25 years (1992-2016) and recorded the necessary data in a Microsoft Excel worksheet in tabular format. Videos for the following two penalty shootouts from UEFA Champions League were not found in the online archive: PSV Eindhoven vs. Olympique Lyonnais, 2004-05 and FC Schalke 04 vs. FC Porto, 2007-08. The dataset includes 470 penalty kicks from 47 shootouts. These tournaments are highly prestigious and competitive because of the achievement standards and financial rewards associated with them. Table 1 presents the distribution of penalty shootouts in recorded data for each of the three tournaments.

Table 1

<table>
<thead>
<tr>
<th>Tournaments</th>
<th>Number of Penalty Shootouts</th>
<th>Percentage Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIFA World Cup</td>
<td>18</td>
<td>38%</td>
</tr>
<tr>
<td>UEFA European Championship</td>
<td>15</td>
<td>32%</td>
</tr>
<tr>
<td>UEFA Champions League</td>
<td>14</td>
<td>30%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>47</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>
Out of 47 penalty shootouts, 10 of them went on to ‘sudden death’ and the ‘sudden death’ data was considered as part of that shootout. Since this study only focuses on whether or not goalkeepers dove in the correct direction or not, the kicks which were shot outside the goal were also included to examine whether the goalkeeper dove in the correct direction or not. The list of variables included and coded in the data table are as follows-

1. **shootout_id**: Each penalty shootout will be given a unique ID in the form of positive integers (1, 2, …, 47)

2. **team_id**: Each team involved in a penalty shootout will be given a unique ID by using the set of positive integers (1, 2, 3, 4 …). These ID’s will be unique at the shootout_id level. For example, if one team appears in more than one penalty shootout, each situation will receive a unique team_id.

3. **kick_id**: Each penalty kick for a particular team will be coded in the order of 1, 2, 3, 4…

4. **kick_dir**: Kick directions for each kick will be recorded as follows- if the kicker shoots to the left side of the goalkeeper, the value of kick_dir will be 0; if he shoots to the right, the value be 1; and if he shoots to the center, the value will be 2. For this study, center will be considered as approximately the region that the goalkeeper could have reached without moving (Palacios-Huerta, 2002; 2003).

5. **dive_dir**: Goalkeepers’ dive directions for each kick will be recorded as follows- if the goalkeeper dives on the left, the value of dive_dir will be 0; if he chooses right, the value be 1; and if he didn’t move or jump (stayed in center), it will be 2.

**Analysis Method**

In order to examine the presence of gambler’s fallacy and hot outcome patterns in
goalkeepers’ choices, the overall number and percentage of times the goalkeepers dove in the opposite and same directions following sequences (varying length) of repeated correct or incorrect predictions in one particular direction were counted. Those kicks where the kickers shot the ball in the center or the goalkeepers remained in the center during a single penalty shootout were assumed to terminate the sequences as such instances are extremely rare and there is no prior evidence of sequential effects on center choices and vice versa (Misirlisoy & Haggard, 2014; Braun & Schmidt, 2015). In the data the kickers shot in the center 9.79% of times, and the goalkeepers stayed in the center 5.53% of the time. Following a streak of at least one correct prediction in a particular direction (left/right), a dive in the opposite direction of the previous dive was classified as a gambler’s fallacy dive whereas a dive in the same direction was regarded as a hot outcome dive. On the other hand, a dive in the same direction as the previous dive following at least one incorrect prediction in a particular direction (left/right) was counted as a gambler’s fallacy dive and a dive in the opposite direction was considered as a hot outcome dive. However, in each occasion the success and failure probabilities for the goalkeepers remain constant as the kick directions are independent draws from a random process.

Table 2

Examples of Gambler’s Fallacy and Hot Outcome Dives Following One Correct or Incorrect Predictions with Any Particular Direction (Left or Right) in the Data

<table>
<thead>
<tr>
<th>Streak length</th>
<th>Scenarios</th>
<th>Previous kick_dir</th>
<th>Previous dive_dir</th>
<th>Prediction Outcome</th>
<th>Subsequent Dive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gambler’ fallacy</td>
</tr>
<tr>
<td>1</td>
<td>I</td>
<td>1</td>
<td>1</td>
<td>Correct</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>0</td>
<td>0</td>
<td>Correct</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>I</td>
<td>1</td>
<td>0</td>
<td>Incorrect</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>0</td>
<td>1</td>
<td>Incorrect</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 3

*Examples of Gambler’s Fallacy and Hot Outcome Dives Following Two Consecutive Correct or Incorrect Predictions in the Same Direction (Left or Right) in the Data*

<table>
<thead>
<tr>
<th>Streak length</th>
<th>Scenarios</th>
<th>Previous kick_dir</th>
<th>Previous dive_dir</th>
<th>Prediction Outcome</th>
<th>Subsequent Dive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>Correct</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>I</td>
<td>1</td>
<td>1</td>
<td>Correct</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>Correct</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>0</td>
<td>0</td>
<td>Correct</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>0</td>
<td>Incorrect</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>I</td>
<td>1</td>
<td>1</td>
<td>Incorrect</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>Incorrect</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>0</td>
<td>1</td>
<td>Incorrect</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4

*Examples of Gambler’s Fallacy and Hot Outcome Dives Following Three Consecutive Correct or Incorrect Predictions in the Same Direction (Left or Right) in the Data*

<table>
<thead>
<tr>
<th>Streak length</th>
<th>Scenarios</th>
<th>Previous kick_dir</th>
<th>Previous dive_dir</th>
<th>Prediction Outcome</th>
<th>Subsequent Dive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>Correct</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>I</td>
<td>1</td>
<td>1</td>
<td>Correct</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>Correct</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>Correct</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>0</td>
<td>0</td>
<td>Correct</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>0</td>
<td>Incorrect</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>I</td>
<td>1</td>
<td>0</td>
<td>Incorrect</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>0</td>
<td>Incorrect</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>0</td>
<td>1</td>
<td>Incorrect</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
<td>Incorrect</td>
<td>0</td>
</tr>
</tbody>
</table>
Tables 2, 3, and 4 illustrate examples of gambler’s fallacy and hot outcome dives following repeated correct or incorrect predictions in the same direction. Proportions of both type of dives were compared against each other after every length of the streak in both situations. Statistically, a binomial test is appropriate to determine the presence of gambler’s fallacy and hot outcome effect (Croson & Sundali, 2005; Braun & Schmidt, 2015) as long as there is no significant overall bias present in the kick directions (i.e., kick directions should be binomially distributed with a probability of 0.5). A binomial test at each level of streak compares the actual dive directions (opposite or same) observed with the baseline hypothesis of 50%. Two-tailed and one-tailed binomial tests were applied to examine whether the goalkeepers were indeed more likely than chance to dive in a particular direction (opposite or same direction) after every streak length (correct or incorrect predictions in the same direction).

To investigate whether the gambler’s fallacy and hot outcome can indeed generate more correct predictions, the correct prediction probabilities of both types of dives (gambler’s fallacy and hot outcome dives) were computed and statistically compared after every streak length. Whenever the goalkeepers dove in the direction of the kick, then it was classified as a correct prediction, otherwise it was counted as an incorrect prediction. If the gambler’s fallacy and the hot outcome are not mere fallacies, then the correct prediction probability for gambler’s fallacy and hot outcome dives should go up progressively as the length of the streak increases and should be significantly higher in comparison to the non-gambler’s fallacy and non-hot outcome dives. That means, as the length of the streak increases, the goalkeepers would become more likely to succeed (make correct predictions) if they dive according to gambler’s fallacy or hot outcome than would be for the opposite strategy. Since gambler’s fallacy and hot outcome are in opposition to each other (i.e., goalkeepers can only dive in the opposite or same direction
following series of dives in the same direction), the non-gambler’s fallacy dives are actually hot outcome dives and non-hot outcome dives are indeed gambler’s fallacy dives. Hence, both of them cannot produce more correct predictions simultaneously (i.e., either one of them will be more effective or both of them will be mere fallacies only). Fisher’s exact test was applied to statistically compare the correct prediction probabilities. This exact test is more accurate than the chi-squared or G-test of homogeneity of proportions when the overall sample size is small (typically less than 1000). The testable null hypothesis for the Fisher’s exact test is as follows:

\[ H_0: p_1 = p_2 \text{ vs. } H_1: p_1 \neq p_2 \]

Where \( p_1 \) and \( p_2 \) are correct prediction probabilities for gambler’s fallacy and hot outcome dives. Acceptance of the null hypothesis at a particular step would indicate that one type of fallacy does not result in significantly higher chances of making a correct prediction in comparison to the other in that step. Xu and Harvey (2014) in their online gambling study used a similar kind of methodology to determine the existence of hot hand and gambler’s fallacy streaks. However, they focused on the streaks of winning and losing, whereas this study deals with streaks of outcomes. This study is based on a modification of their methodology to address the particular foci of this study.
CHAPTER IV. ANALYSIS AND RESULTS

Descriptive Analysis

Among the 405 penalty kicks, after excluding the kicks those went to the center or where goalkeepers remained in the center, 49.6% of the kicks were hit to the left side of the goalkeepers and 50.4% to the opposite direction (i.e., right side) whereas goalkeepers dove 48.1% of occasions to the left and 51.9% of times to the right. To investigate for any overall bias present in the left or right kick and dive directions, two-tailed binomial tests were performed with the baseline hypothesis of 50%. In statistics, a binomial test (an exact test) is used when there are two categories (e.g., coin toss) and the null hypothesis is that those two categories are equally likely to occur. Since both the kick and dive directions in penalty shootouts are i.i.d. random variables (Palacios-Huerta, 2002), a binomial test is appropriate in this context. A binomial test indicated that kickers were no more likely to shoot on the right than that would be expected by chance (i.e., kickers are equally likely to shoot in either direction) (right = .50; exact binomial $p = .921$, two-tailed; $p > 0.1$). The null hypothesis also could not be rejected for the goalkeepers’ dive directions (right = .52; exact binomial $p = .487$, two-tailed; $p > 0.1$) indicating that overall goalkeepers were equally likely to dive in either direction (left or right). Both the findings are consistent with the previous research (Misirlisoy & Haggard, 2014).

Gambler’s Fallacy and Hot Outcome

The analyses began by looking for gambler’s fallacy and hot outcome patterns in goalkeepers’ choices following a streak length of n. A streak length of n is simply the number of times goalkeepers correctly or incorrectly predicted the kicks that were shot in the same direction (left/right).
Table 5 describes the overall goalkeeper data for frequency of dives in the opposite or same direction in the next kick after streaks of correct predictions (in the same direction) of different lengths.

Table 5

*Frequency of Goalkeepers’ Dives in the Opposite or Same Direction in the Next Kick After Streaks (of Correct Predictions in the Same Direction) of Varying Length*

<table>
<thead>
<tr>
<th>Sequence Length</th>
<th>Opposite Direction</th>
<th>Same Direction</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>99</td>
<td>63</td>
<td>162</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>5</td>
<td>29</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

As shown in Table 5, following one correct prediction in any direction (left or right), goalkeepers dove in the opposite direction 99 (61%) times and in the same direction 63 (39%) times. At this point, a dive in the opposite direction would be counted as a gambler’s fallacy dive and a dive in the same direction would be regarded as a hot outcome dive. Of the total 29 dives that took place after a streak length of two (two consecutive correct predictions to the left or right), 24 (83%) were gambler’s fallacy dives and the other 5 (17%) were hot outcome dives. Similarly, after a streak length of three (three consecutive correct predictions to the left or right), the number of gambler’s fallacy and hot outcome dives were 3 (100%) and 0 (0%), respectively.

Figure 2 shows overall goalkeeper data for the percentage of dives in the opposite or same direction in the subsequent kick after streaks of correct predictions (in the same direction) of different lengths.
Figure 2. Percentage of goalkeepers’ dives in the opposite or same direction in the next kick after streaks (of correct predictions in the same direction) of varying lengths.

As can be seen in the Figure 2, the goalkeepers preferred to dive in the opposite direction for the next kick than in the same direction. As the streak length increased, the proportion of dives in the opposite direction increased monotonically. Goalkeepers’ choices therefore show a pattern consistent with gambler’s fallacy. Two tailed and one-tailed binomial tests after every streak length compared the observed proportion of dive directions (opposite or same) in the data with the baseline hypothesis of 50%. The testable null hypothesis after each streak length was-

Two-tailed: $H_0: p = 0.5$ vs. $H_1: p \neq 0.5$

One-tailed: $H_0: p \leq 0.5$ vs. $H_1: p > 0.5$

Where $p$ is the proportion of dives in the opposite direction. The null hypothesis was rejected at the 1% significance level for streak length of one (exact binomial $p = .006$, two-tailed; exact
binomial \( p = .003 \), one-tailed) and two (exact binomial \( p = .001 \), two-tailed; exact binomial \( p = .0005 \), one-tailed). For the streak length of three, although goalkeepers dove in the opposite direction all the times, the null hypothesis could not be rejected at the 10\% significance level (exact binomial \( p = .250 \), two-tailed; exact binomial \( p = .125 \), one-tailed). This is due to the small sampling bias (total number of dives was 3 only), and it is not ideal to comment on the basis of such results. Hence, the streak length of three was not considered for further analysis in this study. This study found statistical evidence for the presence of gambler’s fallacy in goalkeepers’ choices at each streak length. However, there was no evidence for hot outcome effect in the overall goalkeeper data. This is quite obvious as gambler’s fallacy and hot outcome are opposite of each other and cannot be present simultaneously.

Table 6 describes the overall goalkeeper data for the frequency of dives in the opposite or same direction in the subsequent kick after streaks of incorrect predictions (in the same direction) of different lengths.

**Table 6**

*Frequency of Goalkeepers’ Dives in the Opposite or Same Direction in the Next Kick After Streaks (of Incorrect Predictions in the Same Direction) of Varying Length*

<table>
<thead>
<tr>
<th>Sequence Length</th>
<th>Opposite Direction</th>
<th>Same Direction</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52</td>
<td>58</td>
<td>110</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

As presented in Table 6, following one incorrect prediction in any direction (left or right), goalkeepers dove in the opposite direction 52 (47\%) times and in the same direction 58
(53%) times. At this point, a dive in the opposite direction would be counted as a hot outcome dive, and a dive in the same direction would be regarded as a gambler’s fallacy dive. Of the total 17 dives that took place after a streak length of two (two consecutive incorrect predictions to the left or right), 9 (53%) were hot outcome dives and the other 8 (47%) were gambler’s fallacy dives. Similarly, after a streak length of three (three consecutive incorrect predictions to the left or right), the number of gambler’s fallacy and hot outcome dives were 0 (0%) and 2 (100%), respectively. Figure 3 shows overall goalkeeper data for percentage of dives in the opposite or same direction for the subsequent kick after streaks of incorrect predictions (in the same direction) of different lengths.

![Figure 3](image_url)

**Figure 3.** Percentage of goalkeepers’ dives in the opposite or same direction in the next kick after streaks (of incorrect predictions in the same direction) of varying lengths.

As shown in the Figure 3, although goalkeepers slightly preferred to jump in the same
direction for the next kick after a streak length of one, but thereafter they showed a trend to
switch direction. As the streak length increased, the proportion of dives in the opposite direction
increased monotonically, therefore displaying a tendency toward hot outcome behavior.
However, this pattern was not statistically significant. One-tailed and two-tailed binomial tests
after every streak length compared the observed proportion of dive directions (opposite or same)
in the data with the baseline hypothesis of 50%. The testable null hypothesis at each streak
length was –

One-tailed: $H_0: p \geq 0.5$ vs. $H_1: p < 0.5$ (for streak length = 1)

One-tailed: $H_0: p \leq 0.5$ vs. $H_1: p > 0.5$ (for streak length = 2, 3)

Two-tailed: $H_0: p = 0.5$ vs. $H_1: p \neq 0.5$

Where $p$ is the proportion of dives in the opposite direction. The null hypothesis could not be
rejected at the 10% significance level for a streak length of one (exact binomial $p = .317$, one-
tailed; exact binomial $p = .634$, two-tailed), two (exact binomial $p = .500$, one-tailed; exact
binomial $p = 1.000$, two-tailed), and three (exact binomial $p = .250$, one-tailed; exact binomial $p$
= .500, two-tailed). Since the test results for the streak length of three suffers from a small
sampling bias (total number of dives was 2 only), it was not considered for further investigation
in this study. Although goalkeepers showed a pattern toward hot outcome, there was no
statistically significant evidence. There was no evidence for the existence of gambler’s fallacy
either in the overall goalkeeper data.

**Biases or Behaviors**

The analysis began by examining the dives made by the goalkeepers following
sequences of correct predictions (in the same direction) at varying lengths. Figure 4 shows the
overall goalkeeper data for correct prediction probabilities for gambler’s fallacy and hot outcome dives after each streak length. The streak length of three was excluded from the analysis as mentioned earlier because of the small sample size.

Among the 99 goalkeepers who dove in accordance with gambler’s fallacy (i.e., in the opposite direction) for the subsequent kick following one correct prediction (left or right), 56 of them predicted it correctly. The probability of correct prediction was 0.57. On the other hand, 63 goalkeepers dove in the same direction (i.e., in accordance with hot outcome) and the number of correct predictions was 39. Hence, the corresponding correct prediction probability was 0.62. In order to examine whether these probabilities differed significantly or not, Fisher’s exact test was applied. The null hypothesis could not be rejected at the 10% significance level (exact $p = .518$) indicating that there was no significant difference between two correct prediction probabilities.
probabilities.

For goalkeepers who made two consecutive correct predictions in the same direction, the number of gambler’s fallacy and hot outcome dives were 24 and 5, respectively. The corresponding frequency (and probabilities) of correct predictions were found to be 17 (0.71) and 3 (0.60). The two probabilities were not significantly different at the 10% significance level (exact $p = .633$).

The same approach was applied while examining the dives made by the goalkeepers following sequences of incorrect predictions (in the same direction) at varying lengths. Figure 5 shows the overall goalkeeper data for correct prediction probabilities for gambler’s fallacy and hot outcome dives after each streak length. As with the previous analyses, the streak length of three was excluded due to small sample size.

![Figure 5](image)

*Figure 5.* Correct prediction probabilities for gambler’s fallacy and hot outcome dives after each streak (of incorrect predictions in the same direction) length.
Following one incorrect prediction (left or right), 58 goalkeepers dove in accordance with gambler’s fallacy (i.e., in the same direction) for the next kick, and 31 of them predicted it correctly. The probability of correct prediction was 0.53. On the other hand, 52 goalkeepers dove in the opposite direction (i.e., in accordance with hot outcome), and the number of correct predictions was 27. Hence, the probability of correct prediction was 0.52. A two-tailed Fisher’s exact test at the 10% significance level indicated that there is no statistically significant difference between these two probabilities (exact $p = 1.0$).

For goalkeepers who had two repeated incorrect predictions in the same direction. The number of gambler’s fallacy and hot outcome dives for the subsequent kick were 8 and 9, respectively. The corresponding frequency (and probabilities) of correct predictions were found to be 6 (0.75) and 3 (0.33). However, the two probabilities were not significantly different at the 10% significance level (exact $p = .153$).
CHAPTER V. DISCUSSION

This study observed professional goalkeepers’ behavior during the penalty shootouts in order to identify and gather empirical evidence for any potential sequential patterns emerging from their decision-making styles. More specifically, the objective was to examine the presence and influence of the two cognitive biases or statistical illusions in their sequential choices: the gambler’s fallacy and hot outcome. Evidence was found in favor of the gambler’s fallacy, but not for hot outcome when analyzing the dives of goalkeepers who correctly predicted one (or more) kicks in the same direction repeatedly. The analysis suggests that, following one (or more) repeated correct predictions in the same direction, goalkeepers become increasingly more likely to dive in the opposite direction for the subsequent kick from a different kicker than would be expected by chance even though the kick directions are i.i.d draws from a random process. Hence, the goalkeepers dove in accordance with the gambler’s fallacy. Our results are consistent with the findings of previous lab experiments and empirical studies demonstrating negative recency in people’s expectations. Like the bettors in the lottery and dog and horse racing studies (e.g., Clotfelter and Cook, 1993, Metzger, 1985; Terrell, 1998), the goalkeepers also underestimated the probability of a repeat correct prediction in the same direction of the last correctly predicted direction. Such a behavioral pattern is also in agreement with the models of Mullainathan (2002) and Rabin (2002) whereby people overestimate the degree of representativeness of short sequences and expect those sequences to demonstrate “systematic reversals” (Rabin & Vayanos, 2010). Ayton and Fischer (2004) found that people believed in the gamblers’ fallacy for natural events over which they had no control. Goalkeepers’ in this study displayed behavior in accordance with gamblers’ fallacy while predicting the kick directions. This may indicate that they did not believe that kick directions were under their control. Previous
studies on penalty shootouts (Braun & Schmidt, 2015; Misirlisoy & Haggard, 2014) considered all the previous successive kicks in the same direction in general to analyze the existence of gambler’s fallacy behavior in goalkeepers for the subsequent kick whereas this study was more specific when considering the previous successive kicks. Unlike the findings of Braun and Schmidt (2015) who rejected the conclusions of Misirlisoy and Haggard (2014) on the grounds of improper methodology, the results of this study showed serial dependence in goalkeepers’ choices. Although this finding does not directly contradict the results of Braun and Schmidt (or support the findings of Misirlisoy and Haggard), but it adds value to the behavioral literature by more discretely and deeply analyzing the goalkeepers’ choices. Palacios-Huerta (2002) did not specifically consider gambler’s fallacy while analyzing the behavior of goalkeepers during the penalty shootouts but looked for serial dependence in their overall choices (e.g., the streak events were not examined independently). Although the result of the current study does not contradict the results of Palacios-Huerta, it is more in line with the findings of Walker and Wooders (2001) and Kovash and Levitt (2009). That is, professional players systematically deviated from minimax play during competition and gambler’s fallacy was an influencing factor behind such behavior. However, there was no evidence for either gambler’s fallacy or hot outcome when analyzing the dives of goalkeepers who incorrectly predicted one (or more) kicks in the same direction repeatedly. Although goalkeepers demonstrated a tendency toward shifting direction as the length of the streak increased, there was no significant hot outcome effect. This suggests that goalkeepers’ behavioral patterns are in accordance with MSNE in such situations as found in previous game theoretic studies (e.g., Palacios-Huerta, 2002) which looked at their overall behavior. Also, there was no support for goalkeepers’ individual differences in interpreting randomness as this would have only been possible if goalkeepers showed patterns of gambler’s
fallacy in one situation and hot outcome in another situation. Overall, during penalty shootouts professional goalkeepers did not exhibit patterns of positive recency in their expectations but showed negative serial dependence when making decisions on dive directions following a sequence of correct predictions in a particular direction.

Another objective of this study was to examine from the perspective of the adaptive thinking approach whether gambler’s fallacy and hot outcome type predictions would generate more correct predictions from goalkeepers (i.e., whether they are mere fallacies or not). Empirical evidence from the overall data of the goalkeepers suggests that although gambler’s fallacy type predictions showed a monotonically increasing trend in correct prediction rates, there was no statistically significant difference from the opposite (hot outcome) type predictions. That means following a progressively increasing sequence of correct or incorrect predictions in a particular direction, a gambler’s fallacy dive does not ensure statistically increasingly higher chance of success (making a correct prediction) than a hot outcome dive and vice versa. Hence, belief in such statistical illusions do not produce significantly more correct kick direction predictions. Both gambler’s fallacy and hot outcome are mere fallacies and cannot be used adaptively under the adaptive thinking approach. This result is an important finding but warrants further investigation given the paucity of research. Belief in the hot hand has been shown to be a strong cognitive predictive strategy or a fast and frugal heuristic for professional athletes (e.g., Burns, 2001; Arkes 2010; Raab, Gula, & Gigerenzer, 2012; Livingston, 2012), but the same cannot be inferred about gambler’s fallacy and hot outcome.

The findings of the current study indicate that goalkeepers base their decision (consciously or subconsciously) on gambler’s fallacy type prediction while predicting the kick directions following repeated correct predictions with the same direction despite no evidence of
achieving significantly better outcomes (correct predictions). Hence, goalkeepers should refrain from continually adopting such a strategy and instead randomize between the two available strategies which would be in accordance with MSNE. By adopting a gambler’s fallacy strategy, goalkeepers also risk themselves of becoming predictable to penalty kickers. According to the cognitive hierarchy model of Camerer, Ho, and Chong (2004), “players try to anticipate opponents’ decisions but fail to take into account the possibility that their opponents may also be doing this as well as, or even better, than they are” (Misirlisoy & Haggard, 2014, p. 1920). Thus, the kickers could easily monitor for sequential regularities of the goalkeepers’ behavior and could exploit it. On the other hand, the goalkeepers did not exhibit any such patterns while predicting the kick directions following a series of incorrect predictions with the same direction, and none of the fallacies were significantly more effective than the other. In these situations, goalkeepers apparently randomize between the available strategies rather than utilizing any irrational sequential pattern, and this behavior is in fact the optimal strategy as per the previous game theoretical studies.

There are some limitations to this study. In gambler’s fallacy and hot outcome, it is assumed that increasing streaks of events influence people’s decisions (Misirlisoy & Haggard, 2014). However, there may be other factors which can also cause such serial dependence, such as the natural side of the kickers, kicking order, and motivational significance. Future studies should take into account these factors while examining gambler’s fallacy and hot outcome patterns in goalkeepers’ actions and determine the independent contribution of the actual sequential effect above all other factors. Also, since this study has found that goalkeepers dive in accordance with gambler’s fallacy following repeated correct predictions in a particular direction whereas no such effect was observed when they had repeated incorrect predictions previously,
future researchers may wish to investigate the causal relationship between the outcomes of the previous repeated predictions (i.e., correct or incorrect) and subsequent gambler’s fallacy type behavior. A second limitation of this research was the relatively small sample size. Previous empirical studies have typically used very large samples in order to investigate the existence of these two cognitive biases and also to examine whether biases can be adaptive or not. This study could not statistically determine their existence and impact beyond the streak length of two. Also, this study did not consider other factors, such as individual skills (e.g., visual skills) or use of other strategies (e.g., lining up off-center) which may also contribute toward correct predictions. Thus, future research should consider a more robust method which can calculate the independent contribution of gambler’s fallacy and hot outcome after factoring out other extraneous variables.

In conclusion, although this study could not establish the adaptiveness of gambler’s fallacy and hot outcome as an effective fast and frugal heuristic for the goalkeepers to anticipate kick directions during penalty shootouts, the finding of the influence of gambler’s fallacy in goalkeepers’ real-world choices and actions reveals the presence of cognitive bias or a systematic pattern of deviation from rationality in their decision-making process, and this finding adds to the existing literature in cognitive and behavioral sport psychology.
REFERENCES


Keren, G., & Lewis, C. (1994). The two fallacies of gamblers: Type I and Type II. *Organizational Behavior and Human Decision Processes, 60*, 75–89.


*Journal of Risk and Uncertainty, 17*, 151–166.


