THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF SUPPLY CHAIN AND INFORMATION SYSTEMS

Heat Check: Holistic Performance and the Hot Hand Fallacy

SEAN DIEHL
SPRING 2022

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Management Information Systems
with honors in Management Information Systems

Reviewed and approved* by the following:

Robert Novack
Associate Professor of Supply Chain Management
Thesis Supervisor

John C. Spychalski
Professor Emeritus of Supply Chain Management
Honors Adviser

* Electronic approvals are on file.

ABSTRACT

This paper aims to reframe the discussion and analysis of the “Hot Hand Fallacy” by incorporating performance psychology into the background assumptions used for conditional analysis of a made field goal. Typical research into the “Hot Hand Fallacy” performs streak analysis and conditional probability analysis on one metric only: field goals. By incorporating the psychological concept of the flow state and broadening the evaluation of a player’s performance to encompass a more holistic depiction of their “hotness” or “coldness,” this paper performs conditional probability analysis on league-wide field goal percentages. The study found that players who outperform their expected statistics in points, rebounds, assists, steals, and blocks at a level higher than league average can be expected to have a higher field goal percentage on their next attempt. The study also found that through only the use of a holistic metric of past performance relative to expected performance, a binary classification model sees a significant increase in accuracy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Dr. Robert Novack, Ms. Debra Rodgers, and Ms. Melissa Paloskey for their support throughout this process. I would like to thank my parents for their support throughout my academic career.

Chapter 1

Introduction of Topic

The "Hot Hand Fallacy" refers to the belief that a basketball player who, having just made a field goal attempt, will see an increase in the likelihood of success in the subsequent attempt. This paper will perform an analysis that reframes the question posed by the "Hot Hand," by considering an alternative question, "Is a player who holistically outperforms their expected averages more likely to make their next field goal attempt?"

The Origin of the "Hot Hand Fallacy"

The "Hot Hand" in basketball is the stuff of legend for players, coaches, fans, and even statisticians. The fallacious nature of the "Hot Hand" was first introduced in 1985 via the watershed study published by Gilovich, Tversky, and Valone, "The Hot Hand in Basketball: On the Misperception of Random Sequences." Ever since, it has been the subject of impassioned debate and meticulous analysis. Basketball players and coaches have demonstrated that they firmly believe in the existence of the "Hot Hand;" they often describe a sensation of "knowing" that the shot will go in before they even release the ball. Statisticians disagree, supported by the initial findings of Gilovich, Tversky, and Valone. Their paper describes the widespread belief in the "Hot Hand" as being a misguided acceptance of the "Law of Small Numbers" – the notion that small samples should represent the larger population from which they are drawn and reflect the population's generative process. Researchers have used statistical analysis to uncover the true nature of the “Hot Hand” – almost always looking for an answer to the same question, “If a player makes a few shots in a row, are they more likely to make the subsequent attempt?”

The “Hot Hand Fallacy” vs. The “Gambler’s Fallacy”

The fallacious nature of the “Hot Hand” is useful to describe misinterpretations of truly independent and identically distributed variables, like the roll of dice or drawing a card from a deck. Hence, it has also earned the name, “Gambler’s Fallacy,” a name which this paper believes to be more appropriate. The “Gambler’s Fallacy” aptly describes the same misconception of the “Law of Small Numbers” as it applies to casino games. For example, a bad gambler might think, “The roulette wheel can’t land on number 13 again – it’s happened three times in a row!” Unfortunately for our fictitious risk-taker, each spin of the roulette wheel is, in fact, independent and identically distributed – the previous outcomes have no influence on the current outcome. Can the same be said for an intrinsically psychologically-influenced endeavor like athletic performance? Does previous experience throughout the course of the competition truly have no impact on the players’ future performance? In previous literature regarding the “Hot Hand Fallacy,” the factors that influence a player’s shooting percentage are based solely on their previous shot attempts. Here, a field goal attempt is analogous to a spin of the roulette wheel – completely independent of previous attempts. Work published in the field of performance psychology, however, suggests that this framework might not be the most suitable for identifying if players can, in fact, get “Hot.”

Applying The Flow State to “Hot Hand” Analysis

In the field of performance psychology, having a “Hot Hand” is referred to as being in a state of “flow.” The concept of flow was first introduced by psychologist Mihaly Csikszentmihalyi in 1975 - ten years prior to the statistical analysis done by Gilovich, Tversky, and Valone - and since then the concept has been equally studied and analyzed in its respective field. Entering a state of flow has been characterized as requiring at least these three conditions: 1) the activity must have a clear goal, 2) the activity must provide immediate feedback, and 3) a balance between the perceived challenge and one’s perceived ability, i.e., one must be confident that they can achieve the goal (Csikzentmihalyi, Abuhamdeh, and Nakamura, 2005). Basketball, like many other athletic endeavors, clearly meets these requirements. In research both into the “Hot Hand” and research into the flow state, athletes have reported experiencing similar phenomena: a loss of a sense of time, an unconscious sense of their performance, a sense of clarity, a loss of the sense of self, and an overwhelming sense of calm have all been reported.

The field of performance psychology has established beneficial association between reception of positive feedback and the performance of the task at hand while in a state of flow (Swann, 2016). Historically, statistical analysis on the “Hot Hand Fallacy” has, however, considered a made field goal attempt as the only possible source of positive feedback. This necessitates reframing the question behind the “Hot Hand Fallacy” from a focus on streak analysis, like, “Will players who make multiple consecutive field goals be more likely to make their next field goal attempt?” towards a broader definition of positive feedback. As a proxy for an athlete being in a state of flow, this paper will include feedback from results other than field goal attempts into the analysis, including assists, rebounds, steals, and blocks. This process allows for widening the scope of feedback that a basketball player receives while they are on the court.

The “Gambler’s Fallacy,” deals with events uninfluenced by performance psychology, while the “Hot Hand Fallacy” is directly tied to human, athletic performance. To further distinguish between these concepts, this paper will be focused on operating under the assumption that producing positive plays and outperforming the holistic expected average will provide positive feedback and, in turn, make that player more likely to be in a flow state. Does this flow state have an impact on the player’s performance on the court? Does the holistic measurement of players performance on the court against their expected averages lend itself as a useful metric for further analysis? And, more generally, do players who outperform their expected averages see an increase in field goal percentage?

Thesis Statement

As calculated through a holistic metric, players will see an increase in field goal percentage when their actual performance exceeds their expected performance.

Chapter 2

Literature Review

Statistical Analysis of “Hot Hand” Streaks

There has been a great deal of statistical research performed on the analysis of streaks in activities that expect independent and identically distributed variables. The initial study of how this affects the “Hot Hand,” published in 1985 by Gilovich, Tversky, and Valone, “The Hot Hand in Basketball: On the Misperception of Random Sequences” focused solely on the statistical analysis of streaks, or runs, in the makes and misses of players field goal attempts. They decisively state that “The belief in the hot hand and the “detection” of streaks in random sequences is attributed to a general misconception of chance according to which even short random sequences are thought to be highly representative of their generating process” (p. 295). This follows the “Law of Small Numbers,” first established by Tversky and Kahneman in 1971 – a social cognitive bias wherein people misplace their belief in the randomness of streaks. As it relates to basketball, this means that players, fans, and coaches are misinterpreting the randomness that factors into players’ makes and misses forming themselves neatly into extended streaks – getting “hot” for streaks of makes and “cold” for streaks of misses. Their study focused on two essential questions regarding streak analysis: “… the probability of a hit should be greater following a hit than following a miss (i.e., positive association). Second, [the belief in the hot hand implies] that the number of streaks of successive hits or misses should exceed the number produced by a chance process with a constant hit rate (i.e., non-stationary)” (p. 297). These two points of focus, analyzing conditional probability after a make or a miss and analyzing the actual versus the expected number of streaks, have shaped the analysis for the "Hot Hand" for decades to come.

Concerns Regarding Assumptions in "Hot Hand" Literature

The statistical analysis of conditional probability and expected streaks throughout the "Hot Hand Fallacy" rests on key assumptions of independent and identically distributed field goal attempts by each player. Indeed, Gilovich, Tversky, and Valone address this concern in their initial paper, "It may seem unreasonable to compare basketball shooting to coin tossing because a player's chances of hitting a basket are not the same on every shot. Lay-ups are easier than 3-point field goals and slam dunks have a higher hit rate than turnaround jumpers. Nevertheless, the simple binomial model is equivalent to a more complicated process with the following characteristics: Each player has an ensemble of shots that vary in difficulty (depending, for example, on the distance to the basket and on defensive pressure), and each shot is randomly selected from this ensemble. This process provides a more compelling account of the performance of a basketball player, although it produces a shooting record that is indistinguishable from that produced by a simple binomial model in which the probability of a hit is the same on every trial" (p. 297). This major assumption, that each shot is randomly selected from the ensemble of available shots, was disproved in 2014 by Bocskosky, Ezekowitz, and Stein, in their paper, "The Hot Hand: A New Approach to an Old Fallacy." They address the concern that all preceding analysis of the "Hot Hand" has "rested on the assumption that shot selection is independent of player-perceived hot or coldness." This assumption, they proved, was incorrect – and, stood on tenuous ground to begin with, as players have already self-identified as having a belief in the "Hot Hand." Bocskosky, Ezekowitz, and Stein proved that players who have exceeded their expectation over recent field goal attempts can be expected to, on their next attempt, shoot from further away, they can be expected to face tighter defense, and they can be expected to be more likely to take their teams next shot. Each of these factors provides substantial basis for disregarding the initial assumption that players are selecting their shots randomly from the ensemble of available shots. Aharoni and Sarig, in a 2011 paper titled, "Hot hands and equilibrium," also proved that the perceived hotness of a player directly influences the choices that both teams and their coaches make throughout the game. This provides a solid base to substantiate the idea that players and coaches are not acting randomly within each game, but rather constantly adapting their decisions and their performance.

Statistical analysis has disproved other assumptions within the traditional analysis regarding the "Hot Hand Fallacy" as well. In 2018, Muller and Sanjurjo published a paper, "Surprised by the Hot Hand Fallacy? A Truth in the Law of Small Numbers," where they address a substantial bias exists in the conditional dependence of outcomes relative to previous streaks, a concept they dub, "streak selection bias." This analysis does identify each shot as independent and identically distributed and is focused instead on the selection of which shots to consider for conditional probability analysis. They identify that by only measuring specifically after a success, you reduce the number of times in which a success is likely to follow – "for a truly binary event, the expected proportion should be less than 0.5." In an article they published in Scientific American, Miller and Sanjuro state, "in other words, selecting which part of the data to analyze based on information regarding where streaks are located within the data, restricts your choice, and changes the odds". This translates to the idea that Givolich, Tversky, and Valone's original null hypothesis that the expected probability of a hit after a streak of ‘k’ hits minus the expected probability of a streak of ‘k’ misses should equal zero, is, in fact, incorrect.

Regardless of the concerns regarding the statistical assumptions within “Hot Hand” literature and the concerns regarding how the perceived hotness of an individual drives in-game decision-making for players and coaches, there has been surprisingly little research focused on a combination of performance psychology’s flow state and its relation to conditional probability for basketball players. This paper will aim to incorporate a proxy measure for the flow state – one that aligns closely with previous statistical analysis of the “Hot Hand” – outperformance of expected averages in multiple statistical categories.

The Flow State

The concept of a flow state was first introduced by Mihaly Csikszentmihalyi in a 1975 paper, “Beyond Boredom and Anxiety: The Experience of Play in Work and Games.” Csikszentmihalyi has continued to study the concept of flow for decades, influencing a novel field of psychology now known as “performance psychology.” In a book published with the same title, Charles Brown defines performance psychology as “Performance psychology is the systematic application of psychological principles and techniques to performance, particularly when there is a time element and one must perform on demand” (Brown, 2009). The flow state can be characterized qualitatively by an energized focus, a complete intrinsic immersion into an activity. This closely aligns with the interviews performed by Gilovich, Tversky, and Valone with professional basketball players, who described a sensation of “knowing” that a shot would go in after multiple made field goals and a sensation that they “almost can’t miss.”

In 2005, Csikszentmihalyi, Abuhamdeh, and Nakamura published a paper titled “Flow,” where they identify three key characteristics that must be present for an individual to enter a state of flow: 1) the activity must have a clear goals and progress, 2) the task must provide clear and immediate feedback, and 3) good balance is required between the perceived difficulty of the task and the individual’s perceived skills. There have been other attempts at defining what, exactly, the relationship between the individual and the task must be in order to enter a flow state, with as many as seven and eight total criteria being proposed (Schaffer, 2013; Massimini, Csíkszentmihályi and Carli, 1987 - respectively). However, the clear basis for a flow state can be boiled down, so to speak, into the three categories listed above.

The relationship between athletic performance and the flow state has been heavily studied as well, as athletes during competition clearly meet the three provided goals for entering a flow state. Due to the highly qualitative nature of the flow state, research methods such as questionnaires and experience sampling have dominated the literature. Neuroscientific approaches to flow state have been studied in domains outside of sport, ranging from meditation to computer games (Yang, et al, 2019; Harmat, et. al., 2015). However, setting up a neuroscientific approach for studying flow within a competitive athletic environment encounters obvious, as of yet un-surmounted, barriers. Jackson and Marsh published their questionnaire to athletes given in 1998, which features questions such as, “Things just seemed to be happening automatically,” “I had total concentration,” and “I was challenged, but I believed my skills would allow me to meet the challenge,” to which the athletes responded on a scale of 1 – 5 depending on their agreement with the statement. Questionnaires like this one often coincide quite neatly with basketball players’ self-reported feelings regarding being “Hot.” Thus, providing a basis for the assumption that players who self-report as being “Hot” are experiencing a performance within a flow state.

In 2016, Swann published, “Flow in sport,” a holistic analysis of the study of flow and athletic performance. Here, Swann identified the positive association between a state of flow and peak performance, “Flow is highly desirable for athletes due to the association between flow and peak performance.” Competitive athletes themselves have long sought after a psychological advantage over their opponents; it can be assumed that the positive association between the flow state and peak performance can be at the center of this relationship. Tim Grover, a renowned sports psychologist who has worked with top-tier NBA clientele like the late Kobe Bryant, details his process when working with clients in his book, “Relentless: From Good to Great to Unstoppable.” His work with athletes, again, closely matches the ideology surrounding the flow state: he encourages both athletic and psychological training to the point where decisions during performance become automatic. This relationship between automatic performance and peak performance again provides credence to a merger between a statistical analysis of conditional probability when athletes have entered a flow state.

Merging Analysis of Conditional Probability and The Flow State

After reviewing the literature on the statistical analysis of the “Hot Hand” regarding conditional probability and the expected number of streaks – both of which focus entirely on field goal makes and misses – and the literature regarding the flow state and athletic performance, there appears to be a gap. There has yet to be statistical analysis performed on athletes who have entered the flow state. In order to analyze this relationship, this thesis will identify a proxy measure for the flow state that coincides with the qualitative literature: a metric that holistically analyzes a player's actual performance as compared to their expected performance. It stands to reason statistical categories outside of field goals would affect the three qualifications for entering a flow state: rebounds, steals, assists, and blocks all have a clear goal when the task is attempted, all provide immediate feedback relative to the goal of the task, and they all can inspire confidence in the player's ability. By measuring a player's performance relative to their expected performance, this research will identify players who are outperforming expectations to provide a proxy metric for players who are more likely to have entered a flow state during the game.

Chapter 3

Data Methodology

To perform a conditional probability analysis on players who have entered the flow state, this paper will compile a metric, called "Heat," which holistically measures a player's actual performance on the court as compared to their expected performance on the court. The statistical categories that will be taken into account are: Points, Rebounds, Assists, Blocks, and Steals. Their expected performance on the court will be adjusted based on their time in the game, and calculated using the Simple Projection System, which utilizes data from each player's career to provide the basis for the analysis. The actual performance on the court will be taken from play-by-play data for each game in the 2018-2019 season, which tracks each of these categories, including which players for each team are on the court and their cumulative time on the court for each game. Each player's "Heat" metric will then be normalized relative to the league average under- or overperformance for these five statistical categories. Then, to determine the viability of using the "Heat" metric for further analysis, a binary classification model will be trained using only the "Heat" metric and makes and misses. Using only relative past performance during competition to classify shots as makes or misses will provide credence to or disallow further analysis with the metric. Based on each player's "Heat" metric, the data can then be filtered to provide players who are outperforming expectations higher than league average, and underperforming expectations lower than league average, and identify if there is a significant difference between the field goal percentage of the two groups.

Data Collection

Play-By-Play Data

The play-by-play data was collected from eighthirtyfour.com – a site that has compiled the play-by-play statistics based on publicly available data and calculated other useful metrics. The data includes a comprehensive record of each game for the 2001 – 2019 NBA seasons, and the analysis for this paper was performed on the most recent season, the 2018 – 2019 season. Many of the statistics recorded within the data can be disregarded for the purposes of this paper (however, they might be useful in comprising a more comprehensive proxy measure for players who have entered a flow state at some point in the future). The statistical categories relevant for this paper's analysis include: Game ID, Time passed in the game, Home and Away Player ID's for each of the ten players on the court, Home and Away Player ID Play Times for each of the ten players on the court, field goal attempts, free throw attempts, assists, rebounds, steals, blocks, and their respective Player ID associations with who performed each act.

Statistics to track the total cumulative sum of each of the five statistical categories were calculated for all ten players on the floor at any given time. For example, if Player “A” has recorded one assist at the 30 second mark, and another at the 2-minute mark, the assists column for Player “A” would read zero (0) for the time between the start of the game and 29 seconds passed, one (1) for the time between 30 seconds and 1 minute 59 seconds passed, and two (2) at the 2-minute mark forward. Then, these statistics were compiled, by summing them, to provide an unadjusted cumulative performance on the court. For example, if, at any given time during the game, Player “A” had recorded 1 assist, 1 block, 1 steal, 2 points, and 2 rebounds, the cumulative metric would record seven (7). If they were to make another two-point field goal, the cumulative metric would then record nine (9). So, for each player on the floor at any given time, new columns to calculate their cumulative statistics up to that point in the game were calculated. Because the column for field goal makes automatically updates the player’s total points column within the matching associated time, these values were shifted down one row within the dataset. Otherwise, when filtering the data to analyze the conditional probability of field goal percentage given a player’s “Heat,” the total points would inaccurately reflect the player’s points before the field goal in question.

Expected Average Statistics

The expected average statistics were taken from BasketballReference.com, utilizing a method known as the Simple Projection System. The Simple Projection System utilizes a player’s statistics throughout their career to provide a baseline for what they can expect to produce in the upcoming season, so the data for the expected averages statistics was calculated from the data available at the end of the 2017-2018 season. Players without historical data to draw from are awarded expected averages consistent with the median.

The methodology behind the Simple Projection System, is, aptly, quite straightforward:
1) apply a relative weight of 6 to the most recent season, a weight of 3 to the second-most recent season, and a weight of 1 for the third-most recent season. With these weights, calculate a sum of minutes played.
2) The same weights are then applied to a sum for each statistical category.
3) Then, calculate the weighted sum of each statistical category that a league-average player could expect to record if they played the same minutes calculated in step one, then scale this to 1000 minutes.
4) Then, the per 36-minute value can be calculated by taking the sum of steps 1 and 2 and dividing by the sum of step 3 plus 1000 minutes.

The result, then, yields an expected average for the upcoming season per 36-minutes.

Although a seemingly crude metric, it yields surprisingly accurate results: the summed "Heat" metric projections for the 2018-2019 season are plotted against the actual statistics in Figure 1. Figure 1 provides the removal of one outlier observed in the initial regression – a player by the name of Qi Zhou, who recorded a single minute played and made his only field goal attempt, translating into a per 36-minutes average of 72 points. The light blue shaded region surrounding the regression line represents a confidence interval of 90%.

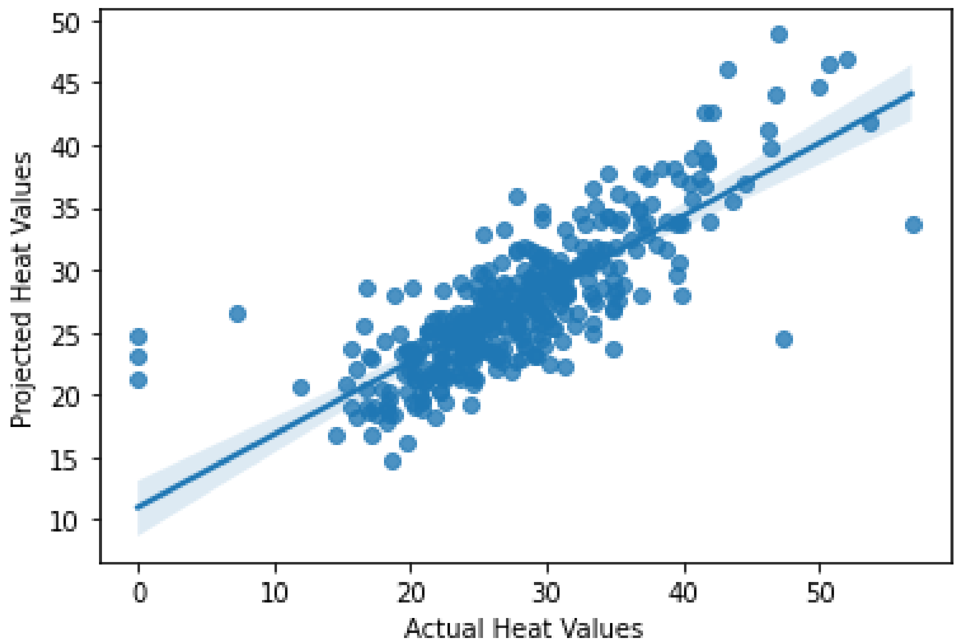

Figure 1 - Actual vs. Projected Heat Values

Although imperfect, the Simple Projection System clearly provides an acceptable baseline from which we can base our expected average statistics, yielding an R-Squared value of 0.645 and a RMSPE of around eleven percent, after removing the aforementioned outlier in Qi Zhou (as well as the three players listed above with total heat values of 0, who did not play in the 2018-2019 season). Given the variability with which players can be expected to perform on a nightly basis, these values were considered acceptable for the purposes of this paper. Indeed, creating a model that more accurately projects individual statistics would constitute its own dedicated research, and thus, falls outside the scope of this paper.

For each of the expected average per 36 statistics, the values were than transformed into a "per-second" basis. This transformation allows for the accurate projection of these statistics relative to the time that each player has played. Multiplying the time played during the game by the expected average per second statistics provides a yardstick from which the actual cumulative performance throughout the game can be measured.

For all ten players on the floor, there are columns tracking their actual cumulative performance and their expected cumulative performance, relative to the time they have played in the game, for each of the following statistical categories: points, rebounds, assists, steals, and blocks.

Heat Metric

Calculating the "Heat" metric was performed by subtracting the corresponding expected statistics from the actual statistics, for every player on the floor in each statistical category. Then, this difference was summed to provide a single, holistic metric for the difference between actual performance and expected performance, relative to time played in the game.

Then, the league mean and league standard deviation for the "Heat" metric was calculated, for use in translating the unadjusted "Heat" metrics into a normalized z-score representation. This allows for the analysis to focus not only on players who are outperforming their expectations but also are doing so at a level higher than the league average. The league mean of the "Heat" metric was 3.07, while the standard deviation was 5.04. This falls in line with expectations already discussed regarding how team behavior changes relative to perceived hot or coldness – players and coaches adjust their strategy based on how hot they perceive a player to be. Thus, players who are severely underperforming can be expected to be sidelined by their coach, resulting in a slightly positive league average for the "Heat" metric. The standard deviation being higher than the league average results in players within the first standard deviation still underperforming their expected averages, which, again, falls in line with what can be expected.

Viability of the Heat Metric

With the normalized “Heat” metrics represented as z-scores relative to the league average, this study will gauge the viability of the “Heat” metric as it pertains to binary classification of field goals into makes and misses. In order to test the viability of the “Heat” metric – that is, to test the viability of past performance relative to expected performance and its impact on players’ field goal percentages – for future analysis, this paper will conduct a basic classification model test. By feeding the data for the “Heat” metric, along with makes and misses, it can be gauged whether or not the “Heat” metric provides significant insight into whether players will make or miss their field goals. In order to test this classification model, a simple, “dummy” classification model will be performed as well. This “dummy” classification model will provide predictions for makes and misses on a stratified basis; it will make predictions based on the distribution of makes and misses within the season’s dataset. For example, if there were 99 makes and 1 miss, it would guess 99 makes and 1 miss on a random basis. Based on historical season field goal percentages, this paper will expect the total field goal percentage to be just under fifty percent. In turn, a stratified prediction model would be correct about fifty percent of the time. Thus, it is not the hope that the ”Heat” metric provides a large boost to the accuracy with which the model can classify field goals as makes or misses, only that it provides a statistically significant difference against the stratified model. This would prove that past performance relative to expected performance can impact the accuracy of a binary classification model for field goals and provide credence to the incorporation of “Heat” into this study’s conditional probability analysis and incorporation into future analysis as well. Using past values to predict future values is not a novel concept with time-series datapoints - the Simple Projection System utilizes the same core ideology. However, the incorporation of past performance relative to expected performance has not yet been tested with a conditional probability analysis of field goal percentage. The viability of the metric will be determined by a hypothesis test regarding two systems of proportions, comparing the “dummy” classification model’s stratified prediction accuracy with the model trained on the “Heat” metric’s accuracy. View Figure 2 and Figure 3 for the formula of a hypothesis test regarding a system of two proportions – note in Figure 3 that the numerator is calculated by summing the number of correct classifications made.

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Figure 2 - Formula for Hypothesis Test Regarding a System of Two Proportions

$$\hat{p} = \frac{Y_1 + Y_2}{n_1 + n_2}$$

Figure 3 - P-hat Value for Hypothesis Test

Conditional Probability Analysis

After performing z-score normalization for the “Heat” metric and determining the viability of the “Heat” metric for its use in further analysis, this study will then analyze the field goal percentages of players in the flow state by performing hypothesis tests concerning two proportions on the field goal percentages of players who are under- and overperforming, with each being compared to the overall league’s field goal percentage. Players who are underperforming will be denoted by a “Heat” metric of less than 0, while players who are overperforming will be denoted by a “Heat” metric of greater than 0.

Data Methodology Summary

After compiling the “Heat” metric for each player on the court, this paper will research the viability of the metric for its use in further analysis. This will be performed through a comparison of the accuracy between two binary classification models predicting field goals as makes or misses – one making stratified predictions and the other trained only on the “Heat” metric. The following provides the null hypothesis and the alternative hypothesis for this significance test.

H0 = There is no significant difference between the accuracy of the two classification models

H1 = The classification model trained on “Heat” is significantly more accurate than the classification model making stratified predictions.

This paper will then perform a conditional probability analysis on field goal percentages of players with a positive “Heat” and a negative “Heat.” This analysis will be performed through a hypothesis test concerning two systems of proportions. The following provides the null hypothesis and the alternative hypothesis for this test.

H0 = There is no significant difference between the field goal percentages of players with a positive “Heat” metric relative to the league’s total field goal percentage

H2 = Players with a positive “Heat” metric have a significantly higher field goal percentage than the league average.

The p-value used for both tests will set at 0.05.

Chapter 4

Data Analysis

The following analysis will focus on the performance of two binary classification models, each of which will be making predictions on whether a field goal will be a make or a miss, and a conditional probability analysis of field goal percentages. One classification model will be making stratified predictions based on the proportional distribution of makes and misses within the dataset, the other will be trained to make predictions on the “Heat” metric. The binary classification model this paper chose to run is a boosted random forest classification model provided by the popular python open-source library XGBoost. As a brief introduction to Gradient Boosting Algorithms, from AWS’ description of XGBoost, “XGBoost minimizes a regularized (L1 and L2) objective function that combines a convex loss function (based on the difference between the predicted and target outputs) and a penalty term for model complexity (in other words, the regression tree functions). The training proceeds iteratively, adding new trees that predict the residuals or errors of prior trees that are then combined with previous trees to make the final prediction. It's called gradient boosting because it uses a gradient descent algorithm to minimize the loss when adding new models.” In other words, the trained model will make decisions about its predictions based on the value of the “Heat” metric – using only past performance relative to expected performance as the basis for its predictions.

Analysis concerning the conditional probability of a made field goal will also be presented. The conditional probability will stem from a player’s “Heat” metric being positive, which can easily be obtained by providing a filter on the dataset. This field goal percentage will then be measured against the league’s total field goal percentage, again using a hypothesis test concerning two systems of proportions.

All the code and data used for this research is publicly available on GitHub at the following URL address: https://github.com/SPDiehl/thesis

Assumptions Within Analysis

Please note some significant assumptions will be made regarding the methodology of this analysis. First, players outperforming expectations at a level higher than league average are assumed to be more likely to be in a flow state. Although certainly an imperfect measure of flow, this assumption coincides closely with the literature on the topic and allows the research performed to bridge the analysis between performance psychology and statistical analysis of the “Hot Hand.” The analysis also assumes that the expected average statistics calculated by the Simple Projection System are accurate to a significant degree, allowing their use for comparison against the actual statistics recorded in the play-by-play data. This research also assumes that the data provided by eighthirtyfour.com delivers significantly accurate play-by-play data – a “sense check” was performed to ensure that random samples of the data matched other publicly available play-by-play data sources. However, this was not comprehensive and did not cross-reference every provided datapoint to other public sources.

Classification of Field Goals as Makes or Misses

Table 1 provides a sample subset of the data after filtering the dataset to provide only the attempted field goals throughout the season and the associated shooting player’s “Heat” score. This data was used for the binary classification model. The “MAKE” column is denoted as a 0 when the player missed the field goal attempt, and as a 1 when the player made their field goal attempt. As this data represents the first field goal attempts for the game, it follows that the “Heat” metric would be slightly negative for each shot attempt, as the players’ expected average statistics calculated on a per second basis could not be expected to be outperformed so early in the game. This, of course, can be expected to return to mean as the time in the game continues.

SHOOTERS_HEAT	**MAKE**
-0.637280	0
-0.643332	0
-0.695586	0
-0.699186	1
-0.699186	0

Table 1 - Sample of Training Data for Classification Model

The number of total shots throughout the 2018-2019 season totaled 219,458: the number of rows in the complete data table. A breakdown of the total field goal percentage can be seen in Table 2.

Total Makes	**Total Misses**	**Total Attempts**	**Total FG %**
101062	118396	219458	46.051%

Table 2 - Total Field Goal Data

When performing stratified predictions on this data, the “dummy” classifier achieved an accuracy of 50.43 percent. This is calculated by the number of True Positives plus True Negatives divided by the Total Number of Predictions. In context of this paper, this is the number of correct predictions of makes and misses divided by the total number of shots.

When the XGBoost classification was trained on the above dataset, it achieved an accuracy of 54.03 percent. In performing a hypothesis test concerning two systems of proportions, the z-score returned provided a value of -22.24 – signifying a significant difference between the classification algorithm’s ability to classify shots as makes or misses, based only on the player’s performance relative to their expected performance. Thus, for our first significance test regarding the accuracy of the classification models, we can reject null hypothesis. The significance test provides the acceptance of the alternative hypothesis H1. The classification model trained on the "Heat" metric was significantly more accurate when making predictions as to whether field goal attempts were makes or misses.

Conditional Probability Analysis

Similarly, the dataset can be filtered to return values for players with a positive and a negative "Heat" metric. Tables 3 and 4 provide the data used for the conditional probability analysis of field goal percentage.

Heat Value	**Performance**	**Makes**	**Misses**	**Attempts**	**FG %**	**Z-Score**
Heat > 0	Overperforming	35448	39790	74238	47.114 %	-5.04964576
Heat < 0	Underperforming	51974	60899	112873	46.046%	0.023461895

Table 3 - Conditional Probability of a Field Goal Based on Heat

The hypothesis test concerning two systems of proportions for each group was performed relative to the data provided in Table 2, that is, relative to the league average. With a z-score of -5.05, the null hypothesis for the conditional probability analysis can be rejected. The alternative hypothesis H2 can be accepted. Players with a positive "Heat" metric can be expected to have significantly higher field goal percentages than the league average. The reverse was also tested – the field goal percentage of players with a negative "Heat" metric as compared to the league average field goal percentage. Interestingly, the reverse is not statistically significant – players underperforming cannot be expected to perform worse than league average. This, however, could be due to the discussed adjustments that coaches and players might make based on a player's perceived hotness or coldness.

Chapter 5

Conclusion

A Holistic Relative Performance Metric and Field Goal Percentage

The results from the classification model and the conditional probability analysis provide a substantial result: by holistically measuring past performance relative to expected performance this research found that players outperforming expectations are more likely to make their next field goal attempt. The “Heat” metric as a measure of holistic performance relative to expected performance also provides a significant increase in accuracy to a binary classification model trained to predict shots as makes or misses. It is difficult to truly incorporate an accurate, real-time, quantitative measure of athletes in the flow state; however, these findings suggest that players who are more likely to have entered the flow state can be expected to be more likely to make their next field goal attempt.

It is difficult to accurately identify in real-time players who have entered the flow state. Using a proxy metric that calculates the difference between players’ actual cumulative performance and their expected cumulative performance throughout the game allows for an imperfect substitution for a quantitative measure of the athletes’ neurological states. The proxy “Heat” metric was determined to be suitable due to the close association between athletes reported feelings during both a flow state and while feeling “Hot,” and the association of peak performance in both. The current calculation of the “Heat” metric can serve as a foundation from which further analysis for past performance relative to expectations can influence a player’s future performance.

It is difficult, as well, to calculate real-time expected cumulative statistics for player’s during competition. Confounding factors such as the competition, the player’s matchups at any given time, the context of the game (a team pushing for the play-offs versus a team that is purposely performing poorly to improve their likelihood of landing the top draft pick), the coach’s strategy, and many others make these cumulative statistics difficult to accurately calculate. The close relationship between the model’s projected per-36 statistics and the actual on-court performance provides a baseline for calculating the “Heat” metric but does not provide a safeguard against these confounding factors.

Using only past performance relative to actual performance, a binary classification model was able to significantly increase its accuracy in predicting shots as being makes or misses. Using the same metric, players who were outperforming their expectations were found to be more likely to make their next field goal attempt as compared to the league average.

Topics to Explore Further

A more holistic measure of conditional probability analysis would be composed of both a more accurate measure for players competing in a flow state and a more accurate model to calculate expected statistics, as well as a player-by-player breakdown of this analysis. There are several topics and questions to consider for further research.

When are players in a state of flow?

This paper assumes that players who have outperformed their expectations are more likely to be performing in a flow state. As the quantitative measures of neurological states are difficult to implement into a competitive athletic environment, it seems that proxy measures are currently the best alternative for quantitatively studying the relationship between the flow state and an athlete's performance. However, implementation of these proxy measures can, and should, vary widely – this paper used past performance relative to expectations due to the association between the flow state and peak performance. Other measures could be an experience sampling survey completed before, during, and after the course of competition or simply a survey relating to the players confidence throughout the game.

How does one measure who gets "Hot"?

This paper measures players who are "Hot" as having outperformed their expectations holistically – that is, it takes other statistics outside of field goals into account. Traditional analysis of the "Hot Hand" has focused solely on field goals; this paper broadened the scope of this definition and asked a different question than conditional probability based on makes and misses. Measuring when and why a player gets "Hot" proposes difficult challenges to future researchers. Still, measuring what players mean when they report feeling "Hot" might be more difficult still. Professional players, experts in their field, continually report similar feelings associated with the traditional "Hot Hand Fallacy." This does not, however, provide a clear path to a quantitative metric to incorporate into predictive analytics.

What is the most accurate way to calculate expected statistics per second?

Calculating expected average statistics on a second-by-second basis poses its own difficulties. Controlling for matchups, coaching decisions, the context of the game, etc., would provide for a more accurate system to calculate expected average statistics. A predictive model that controls for all confounding factors might not ever become publicly available for academic research. Comparing cumulative performance on a second-by-second basis against expected cumulative performance within these guidelines still might not provide an acceptable measure of the flow state or provide accurate insight into when a player feels “Hot”. Even so, a more comprehensive model mapping the relationship between expected average statistics and actual statistics could provide interesting insights into player performance.

Can players get “Hot” in other sports?

While this paper focused on studying the “Hot Hand” in basketball, there are as many potential studies of the “Hot Hand” as there are competitive sports and, more broadly, competitions. Can chess players get “Hot”? Can golfers or baseball players or NASCAR drivers? Expanding the study of the relationship between the flow state and conditional probability of any metric, winning a chess match, scoring a birdie on the next hole, striking out more batters than normal, or winning multiple consecutive races, would provide a further basis for accepting that athletes in the flow state are associated with peak performance. The analysis of streaks and conditional probability would have to be similarly broadened, as in this paper, to provide a holistic proxy measure of the competitor’s psychological state.

BIBLIOGRAPHY

Aharoni, Gil & Sarig, Oded. (2011). Hot hands and equilibrium. Applied Economics - APPL ECON. 44. 1-12. 10.1080/00036846.2011.564141.

AWS. (2022). XGBoost Algorithm. AWS Documentation. https://docs.aws.amazon.com/sagemaker/latest/dg/xgboost.html

Bocskoscky, Ezekowitz, and Stein. (2014). The Hot Hand: A New Approach to an Old 'Fallacy'. Harvard University.

Brown, C. H. (2009). *Performance Psychology*. American Psychological Association.

Davis, M. S. (1977). [Review of Beyond Boredom and Anxiety: The Experience of Play in Work and Games., by M. Csikszentmihalyi]. Contemporary Sociology, 6(2), 197–199. https://doi.org/10.2307/2065805

Davis, M.S., & Csíkszentmihályi, M. (1977). Beyond Boredom and Anxiety: The Experience of Play in Work and Games. Contemporary Sociology, 6, 197.

Gilovich, Thomas; Tversky, A.; Valone, R. (1985). The Hot Hand in Basketball: On the Misperception of Random Sequences. Cognitive Psychology. 17 (3): 295-314.

Grover, T. S., & Wenk, S. (2014). Relentless: From Good to Great to Unstoppable. Simon and Schuster.

Harmat, L., de Manzano, Ö., Theorell, T., Högman, L., Fischer, H., & Ullén, F. (2015). Physiological correlates of the flow experience during computer game playing. International journal of psychophysiology: official journal of the International Organization of Psychophysiology, 97(1), 1–7. https://doi.org/10.1016/j.ijpsycho.2015.05.001

Jackson, S. A. and Marsh, H. W. (1996). Development and Validation of a Scale to Measure Optimal Performance: The Flow State Scale. Journal of Sport and Exercise Psychology. 18. 17-35.

Jackson, S. A., Ford, S. K., Kimiecik, J. C., et al. (1998). Psychological Correlates of Flow in Sport. Journal of Sport and Exercise Psychology. 20: 358 – 378.

Miller, J and Sanjurjo, A. 2018. Surprised by the Hot Hand Fallacy? A Truth in the Law of Small Numbers. https://marketing.wharton.upenn.edu/wp-content/uploads/2018/11/Paper-Joshua-Miller.pdf

Miller, J and Sanjurjo, A. (2018, March 28). *Momentum Isn't Magic—Vindicating the Hot Hand with the Mathematics of Streaks*. Scientific American.
https://www.scientificamerican.com/article/momentum-isnt-magic-vindicating-the-hot-hand-with-the-mathematics-of-streaks/

Oppland, M. (2019, November 20). *8 Ways To Create Flow According to Mihaly Csikszentmihalyi*. Positive Psychology. https://positivepyschology.com/mihaly-csikzentmihalyi-father-of-flow

Saini, Udam and Evans, Katherine. eightthirtyfour.com/data.

Swann, Christian F.. "Flow in sport" (2016). Faculty of Social Sciences - Papers. 2723. https://ro.uow.edu.au/sspapers/2723

Yang, X., Cheng, P., Lin, L, et al. (2018). Can an Integrated System of Electroencephalography and Virtual Reality Further the Understanding of Relationships Between Attention, Meditation, Flow State, and Creativity?. Journal of Educational Computing Research. Vol 57(4) 846 – 876. DOI: 10.1177/0735633118770800 journals.sagepub.com/home/jec

ACADEMIC VITA

EDUCATION

The Pennsylvania State University – Schreyer Honors College University Park, PA
Smeal College of Business | B.S. in Management Information Systems with Honors Distinction | *Spring 2022*
Minor: Spanish

WORK EXPERIENCE

Data Science Intern at Inveniam Capital Partners Virtual Internship
July – Oct 2021

- Created synthetic data representing capital calls and distribution notices to effectively account for industry variation and improve training for the firm's NLP algorithm
- Implemented web-scraping to systemically provide accurate bond purchasing data, categorized by municipality, for Private Equity clients monitoring investments
- Assisted the lead Data Scientist in various other ad-hoc projects through problem-solving and brainstorming sessions; subsequently executed necessary changes or initiated original development via Python

Data Analyst Intern at LTx Virtual Internship
Summer 2020

- Performed database management, data verification, and data transformation via SQL on financial datasets of upwards of 100,000 data points; emphasized scalable, computationally efficient queries and verification processes
- Implemented cursory file maintenance via Python on the training datasets that drive matchmaking recommendations
- Honed the application's UX as a member of the Business Solutions team to increase functional synergies while maintaining a user-friendly format

INVOLVEMENT

Penn State Dance Marathon and Other Volunteer Experience University Park, PA; Philadelphia, PA
Sep 2013 - Present

- Member of THON Alternative Fundraising Committee for Alpha Tau Omega Fraternity, the top fundraising organization for more than 26 years, subcommittee Alumni Outreach and Development, helping to raise more than $322,000 in 2021
- Member of THON Operations Committee, responsible for logistics and organization, set-up, tear-down, and related activities to ensure that events proceed safely
- Tutored and provided extra-curricular learning assistance to grade school students at La Salle Academy in Kensington, Philadelphia multiple times a year

Nittany Data Labs University Park, PA
Aug 2019 – Dec 2021

- Provided a struggling local company with a full-scale business plan and headed company valuation, strategic positioning, competitive analysis, and operational efficiencies sections; incorporated topics learned from classes and involvement in NDL
- Implemented concepts learned about statistics, machine learning, and organizational management into class assignments and personal projects

Study Abroad Experience Florence, Italy; Barcelona, Spain
May 2019; Jan - March 2020

- Studied the history and development of financial accounting through an immersive experience in the field's birthplace
- Delivered presentations and analysis of short-term projects involving historical and current developments in Venetian bookkeeping, with a focus on blockchain-based financial instruments
- Improved Spanish fluency through daily life and Spanish courses while studying in Barcelona, Spain
- Experiential learning was provided by linking out-of-the-classroom activities with academic study; improved problem-solving skillset and understanding of global influence on financial markets

HONORS & INTERESTS

- **Honors:** Schreyer Honors College Academic Scholarship, AP Scholar with Distinction, National Merit Scholarship Commended Student, HOBY Student Leadership Conference
- **Interests:** Puzzles and Crosswords, Surfing, Reading and Writing, College and NBA Basketball, College Lacrosse
- **Skills:** Python with a comprehensive base of Statistics, Visualization, and ML libraries, SQL, Tableau, Proficient in Microsoft Office, VBA, Conversational in Spanish