University of Texas at El Paso
DigitalCommons@UTEP

Open Access Theses & Dissertations

2013-01-01

Individual Predictors of Hindsight Bias: A Longitudinal Study

Sonya M. Stokes
University of Texas at El Paso, smstokes1707@gmail.com

Follow this and additional works at: https://digitalcommons.utep.edu/open_etd

 Part of the Psychology Commons

Recommended Citation
Stokes, Sonya M., "Individual Predictors of Hindsight Bias: A Longitudinal Study" (2013). *Open Access Theses & Dissertations*. 1938.
https://digitalcommons.utep.edu/open_etd/1938

This is brought to you for free and open access by DigitalCommons@UTEP. It has been accepted for inclusion in Open Access Theses & Dissertations by an authorized administrator of DigitalCommons@UTEP. For more information, please contact lweber@utep.edu.

INDIVIDUAL PREDICTORS OF HINDSIGHT BIAS:

A LONGITUDINAL STUDY

SONYA MARIE STOKES

Department of Psychology

APPROVED:

_____
Osvaldo F. Morera, Ph.D., Chair

_____
Gary P. Braun, Ph.D.

_____
Matthew T. Carlson, Ph.D.

_____
Wendy S. Francis, Ph.D.

_____
Benjamin C. Flores, Ph.D.
Dean of the Graduate School

Copyright ©

by

Sonya Marie Stokes

2013

Dedication

For my husband Brian-
your confidence and unerring faith in me
taught me to believe in myself.

For my parents-
it is only through your sacrifice and support
that I have been able to conquer my ambitions.

INDIVIDUAL PREDICTORS OF HINDSIGHT BIAS: A LONGITUDINAL STUDY

by

SONYA MARIE STOKES, B.A.

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS

Department of Psychology

THE UNIVERSITY OF TEXAS AT EL PASO

August 2013

Acknowledgements

I would like to thank my advisor, Dr. Osvaldo F. Morera, for his support and guidance for the past two years, as well as his encouragement to pursue my master's degree in a year. I would also like to thank my committee members, Dr. Matthew T. Carlson, Dr. Wendy S. Francis, and Dr. Gary P. Braun for their insight and wisdom.

I would also like to thank my colleagues, Dr. Katherine R. White, Dr.Nishad Jabeen, Luis Rangel, Mosi Dane'el, Ciara Kidder, Brandt Smith, and Jessica Shenberger for their encouragement, support, and assistance.

Additionally, I would like to thank my husband, Brian, who, although he passed away too early, had such a profound faith in me that he helped me to believe in myself. Finally, I would like to thank my parents, Fred and Sandy Carpenter, for their unending support and encouragement.

Abstract

The hindsight bias, a person's tendency to overestimate their ability to predict the outcome of an event after the fact, is a phenomenon present in nearly every area of our lives (Arkes, Faust, Guilmette & Hart, 1988). The bias can be mitigated by factors such as analytical and deep thought regarding an outcome and increased expertise (Gray et al., 2007; Knoll, 2009; Arkes et al., 1988). In this study, we examined the growth of hindsight bias over 4 time points using a within-subjects memory model. For our analysis, we quantified hindsight bias in two ways: the number of "flips" a participant committed and the proportion of change in their confidence in their decision (measured on a scale of 0-100). We additionally assessed the relationship between hindsight bias and individual differences such as gender, expertise, self-presentational concerns, need for cognition, and impulsivity and carelessness in social problem solving. We found that participants did not flip a significant amount of their predictions for Week 8 (ß = -.95, SE = .68, ns) but committed 12% more hindsight flips for each successive week (ß = -1.99, .34, p < .001). Participant committed confidence hindsight bias on Game 1 of Week 8 (β = .50, SE = .25, p < .05; 25% of bias possible), Week 9 (β = .889, SE = .39, p < .001; 60% of bias possible), and Week 11, β = .34, SE = .17, p < .05, 13% of bias possible. Additionally, the amount of bias committed on Week 9 reduced by 4% for each progressive game, ß = - 0.08, SE = .04, p < .05. Individual differences were unable to significantly predict variance in the hindsight bias.

Table of Contents

List of Tables

List of Figures

Introduction

In August of 2005, Hurricane Katrina devastated southern Louisiana as it blasted to shore from the Gulf of Mexico. Americans watched in horror as levees were breached and the densely populated areas of New Orleans were flooded, killing over 1800 people and leaving 250,000 more homeless and destitute (Graumann, Houston, Lawrimore, Levinson, Lott, McCown, Stephens and Wuertz, 2005). As the extent of the devastation in the aftermath of the storm became apparent, the country began to question if the tragedy could have been prevented, or at the very least minimized.

Reporters produced documents that had warned the White House of the inevitable catastrophe due to a lack of preparedness (Associated Press, 2006). Hurricane experts lambasted the Army Corps of engineers responsible for designing and constructing the levees, saying they should have recognized the design flaws (Schwartz, 2006). The media even criticized victims for not heeding the flood warnings as soon as they were issued (Forsloff, 2009). Research findings suggest that these accusations may be the result of a phenomenon referred to as hindsight bias, which is the tendency of individuals to overestimate their ability to predict an event after it has occurred.

Hindsight bias was first introduced in 1975 by Dr. Baruch Fischoff, who observed that participants in a study tended to inflate their ability to predict an event after they were informed of the actual outcome. Fischoff (1975) found that when participants were asked to recall their predictions following the event's outcome, their memories of their predictions tended to favor the actual outcome. There are two major categories of hindsight bias: overestimates of prediction accuracy after the fact; and distorted representations of knowledge (Blank, Musch, and Pohl, 2007).

Hindsight bias is pervasive in our daily lives. It exists in the form of the innocuous "I told you so" as well as in weightier circumstances such as courtroom decisions and medical opinions (Arkes, Wortmann, Saville, & Harkness, 1981; Arkes, Faust, Guilmet & Hart, 1988; Blank et al, 2007; Harley, 2007). The phenomenon that Fischoff (1975) first observed nearly 40 years ago continues to have powerful effects on our individual lives as well as society at large, making it an important subject for ongoing study.

Experimental Models of the Hindsight Bias

Studies of hindsight bias take the form of two major experimental models: hypothetical models and memory models. The hypothetical model is a between-subjects design in which an experimental group is given a description of an event and outcome. The participants in the hindsight condition are then asked to disregard the outcome and estimate how likely they would have been to provide correct predictions prior to the event. Participants in the foresight condition provide predictions of the event prior to the event taking place. The foresight condition is given the same event description but do not have knowledge of the outcome of the event. The difference in the accuracy of predictions between the foresight condition and the hindsight condition provides a measure of the hindsight bias. In his pioneering study in 1975, Fischoff employed the hypothetical model when he asked participants to estimate their ability to recall details of historical events. When participants were primed with the actual information, they were more likely to estimate that they could have supplied the information than a control group in a foresight condition (Fischoff, 1975).

In contrast to the hypothetical model, a memory model is a within-subjects design using a pretest-posttest format. In this model, there is only one group of participants and they make a

prediction about an event. At a later point in time, they learn of the outcome of the event and are asked to recall their original prediction.

To date, very few studies have studied the reliability of the memory design and hypothetical design using the same material (Musch and Wagner, 2007). However, Musch (2003) compared the relative effectiveness of the hypothetical and memory models (Musch, 2003). In this study, participants were asked to predict their ability to supply the answers for almanac style questions. Half of the questions were designed as a memory model, in which participants first made the predictions, then were supplied with the correct answers and asked to recall their original predictions (Musch, 2003). The remaining questions were written in a hypothetical model. For these questions, half of the participants were given the almanac question, the answer, and then asked to predict their ability to supply the correct answer (Musch, 2003). Of the remaining participants who acted as controls, half were asked to estimate how likely *they* would have been to supply the correct answer (hypothetical within-subject, WS) and half were asked how likely *others* would have been to correctly answer the question (hypothetical between-subject, BS) (Musch, 2003). Additionally, Musch (2003) measured 10 individual difference variables including social desirability, impression management, dogmatism, and need for cognition.

The personality predictors accounted for more variance in the WS hypothetical model (24%) and the BS hypothetical model (39%) than in the memory model (19%) (Musch, 2003). Musch (2003) also determined the reliability of each model by first dividing questions into groups based on even or odd numbers. After computing a measure of hindsight bias for both groups, they were compared in a correlation. Musch (2003) found that the computed reliabilities of measures of the hindsight bias from the BS hypothetical model ($\alpha = .77$) and a measure of the

hindsight bias from the WS hypothetical model (α = .58) were higher than that of the memory model (α = .49). Additionally, although the hindsight bias was observed in all three conditions, the effects were more prominent in the within-subject (M = 1.81, SD = 2.71) and between subject (M = 1.80, SD = 2.52) hypothetical models than in the memory models, M = 0.69, SD = 1.37, $F(1, 74) = 13.81$, $p < .001$ (Musch, 2003).

However, Musch's study was small with only 80 participants. In addition, his findings have failed to be replicated in other studies. A meta-analysis of 95 studies compared the effect sizes associated with the hindsight bias for the memory and hypothetical designs (Guilbalt, Bryant, Brockway & Posavac, 2004). Although researchers hypothesized that the effect size of studies utilizing a hypothetical design would be larger, the meta-analysis revealed found no significant difference between the two designs (Guilbalt et al, 2004). To our knowledge, there are no additional studies comparing hypothetical and memory models of hindsight bias.

Individual Differences Explanation of Hindsight Bias

Studies of hindsight bias to date have primarily focused on predicting hindsight bias through individual differences and understanding its underlying cognitive processes. For the purposes of this study, we focus on individual differences as predictors. This approach seeks to model the relationship between the unique attributes of an individual and their tendency to display hindsight bias.

Musch and Wagner (2007) studied individual predictors of hindsight bias extensively. They found that self-presentational concern, suggestibility, field dependence, and need for predictability are positively correlated with the display of hindsight bias (Musch & Wagner, 2007). Additionally, intelligence, conscientiousness, and need for cognition all are theoretically negatively correlated to hindsight bias (Musch & Wagner, 2007). However, they

also point out that, to date, evidence of these connections is limited by small sample sizes and inadequate measures of the constructs (Musch & Wagner, 2007).

Need for cognition describes an individual's preference for deep thinking and analysis (Cacioppo, Petty & Feng Cao, 1984). Early research by Slovic and Fischoff (1977) found that asking hindsight participants to produce reasons why multiple experimental outcomes might have occurred is associated with reducing the hindsight bias, leading to the hypothesis that need for cognition might be negatively correlated with hindsight bias. In one study, Veplanken and Pieters (1988) tested this hypothesis when they surveyed 212 citizens in the Netherlands 2 months before and 5 months after the Chernobyl accident. The first survey was designed to assess citizen's perception of risk regarding the use of coal and nuclear energy to generate electricity. One item asked participants if they believed that nuclear energy could lead to a large scale catastrophe. When Chernobyl, a nuclear power plant roughly 900 miles from the Dutch border, exploded, the radioactivity level in the area that had been polled was substantially raised. As a result, the Dutch government expressed concern for livestock and regarding the safety of crops that had been exposed. The impact was felt nationwide.

Five months following the explosion, researchers sent citizens another survey. The survey included a shortened Dutch Need for Cognition Scale (Pieters, Verplanken and Modde, 1987) as well as an item that asked participants how likely they estimated a nuclear disaster to be prior to the explosion. Verplanken and Pieters (1988) found that individuals with low to moderate measures of need for cognition actually displayed significant reverse hindsight bias, which occurs when one recalls lower confidence than originally predicted, while individuals with high need for cognition did not show any significant difference in their hindsight and foresight estimates. The study provides evidence that need for cognition and the display of hindsight bias

are negatively related. Additionally, the study provides a logical basis for hypothesizing that impulsivity or carelessness in problem solving, essentially the absence of analysis and deep thought, might predicts hindsight bias. Recent research conducted at UTEP using a memory model found that increased impulsivity was associated with increased initial amounts of the hindsight bias (Morera, Stokes & Duran, 2013).

Self-presentational concerns are often described in two categories. The first is the desire of individuals to control the way they are viewed by others, a subcategory termed "impression management" (Leary & Kowalski, 1988). A second subcategory titled "self-deception" refers individual's attempt to control their own view of themselves (Musch & Wagner, 2007; Stöber, 2001). However, there is some controversy over whether self-deception is in fact a self-presentational concern or whether it has different underlying causes and should be studied under a separate title (Leary & Kowalski, 1988). High levels of self-presentational concern often affect the individual's judgment and decision making (Leary & Kowalski, 1988). For example, in one survey of 101 college freshman, researchers found that 75% of participants had engaged in at least one risky behavior (such as impaired driving, smoking, and drug use) during their first semester in order to manage their impressions on others (Martin & Leary, 2000). Moreover, researchers found that the exhibition of risky behavior was significantly correlated with situational self-presentational concerns, $r = .34$, $p < .01$ (Martin & Leary, 2000).

The tendency of self-presentational concerns to affect judgment and decision making also seems to extend to hindsight bias (Leary, 1981; Verplanken & Pieters, 1988; Musch & Wagner, 2007). In a study by Musch and Wagner (2007), 632 participants were asked to rate their level of expertise regarding chocolate (2007). Next they were given a piece of chocolate and asked to estimate its cocoa content. They were provided with an anchor value ranging from 10%-90%

and asked how much cocoa they would have estimated was contained in the chocolate if they had not received the value. Not only did participants demonstrate hindsight bias and align their responses with the anchor value they received, but Musch and Wagner (2007) also found that the "experts" tended to exhibit more hindsight bias than "non-experts" as evidenced by the slope of each group's standardized regression line, .39 vs. .20, $p < .05$. The researchers attributed this difference to the desire of "experts" to engage in impression management and appear more knowledgeable (Musch & Wagner, 2007). This finding also suggests that self-presentational concerns might interact with expertise as a predictor of hindsight bias.

Shanteau (1992), a researcher in the field of expert decision making, notes that experts tend to display concerns for their self-presentation, regardless of the domain being observed. Musch and Wagner (2007) theorized on relationship between expertise and self-presentational concern, pointing out that those more familiar with the subject matter may be more inclined to display hindsight bias due to motives to maintain their image as knowledgeable in the area of question.

The relationship between expertise and hindsight bias is unclear. In the study by Musch and Wagner (2007), experts were defined by their self-assessment of their expertise. However, doubt has been cast on whether such a simple measure can truly define and expert. Shanteau (1988) identified 6 characteristics that distinguish an expert: advanced perceptual and attentional abilities, an ability to break down complex problems into simpler parts, creativity in the face of new challenges, self-confidence in the domain of their expertise, an advanced ability to communicate their expertise, and an advanced level of content knowledge (also termed familiarity). According to this definition, it is likely that within the field of hindsight bias most

measures of expertise, including that by Musch and Wagner (2007), more resemble measures of familiarity with the subject matter.

Marks Knoll (2009) also sought to explain the effect of expertise on hindsight bias. However, her measure was composed of trivia questions about the study's subject matter, baseball (Marks Knoll, 2009). Shanteau, Weiss, Thomas & Pounds (2003) argue that such measures do not meet the criteria for establishing expertise since a true expert requires more than just factual knowledge. However, Marks Knoll's (2009) sample included baseball players from her university's team, making it possible that although her measure was really indicative of familiarity, her sample truly included experts.

Marks Knoll (2009) found a U-shaped relationship between expertise and the hindsight bias, where extremely low levels of expertise and extremely high levels of expertise corresponded with low levels of hindsight bias and intermediate levels of expertise showing higher levels of the hindsight bias. The difference in this finding from that of the chocolate study by Musch and Wagner (2007) might lie in the true nature of an expert. Musch and Wagner (2007) were examining familiarity rather than expertise. However, since familiarity is a component of expertise, it is possible that the positive relationship that they observed in their study is analogous to the low to medium range of expertise observed by Marks Knoll (2009; Musch & Wagner, 2007). In Marks Knoll's (2009) study, the participants with higher levels of expertise (the baseball players) possessed other elements of Shanteau's definition of expertise in addition to familiarity. This could indicate that as familiarity alone increases, so does the hindsight bias, but that as the other components of expertise increase, the hindsight bias decreases.

Hindsight Bias Over Time

One area of hindsight bias research that has been largely unexamined is its growth of over time. Many studies have attempted to predict hindsight bias by the time elapsed between pretest and posttest and event and posttest, however a meta-analysis found that overall, findings do not show elapsed time as a significant predictor (Guilbault, Bryant, Brockway & Posavac, 2004). These studies only assessed the recall of a single event. However, in reality, individuals are often repeatedly exposed to events which are similar in nature.

For example, returning to our discussion of Hurricane Katrina, New Orleans is situated on the Gulf of Mexico and is often affected by hurricanes. Between the years of 1851-2004, 273 hurricanes made landfall on the east coast and New Orleans was affected by 49 of those (about 18%; Blake, Rappaport, Mayfield & Landsea, 2005). In a city where hurricane and tropical storm warnings are a relatively common occurrence, does the repetitive nature of the warnings have an effect on the hindsight bias displayed? Does past experience with hurricanes affect the hindsight bias individuals display in subsequent events, especially given the evidence of a relationship between familiarity and hindsight bias?

In August of 2012, almost exactly 7 years after the devastation caused by Hurricane Katrina, Hurricane Isaac caused severe flooding in the areas surrounding New Orleans. Storm surges reached 11 feet in some areas and over 600,000 residents lost electricity, with 400,000 residents still without power nearly a month following the storm (Staff, 2012). However, despite the devastation caused, there was a surprising lack of finger-pointing in the aftermath of the storm. In fact, one editorial in the local paper, the Times-Picayune, lauded the efforts of the government, stating that "it's worth taking a few minutes to praise the government…Yes, some

things went wrong this time around, which is, sadly, bound to happen in a disaster" (Grace, 2012).

The change in response among New Orleansians and the rest of the country despite the shockingly similar circumstances to 7 years prior indicates that perhaps their experiences during Hurricane Katrina and their continued exposure to the threat of natural disasters in the ensuing years affected their perceptions of the adequacy of government response to Hurricane Isaac. Whether hindsight bias changes with repeated exposures, particularly as a function of changes in individual differences, merits further investigation.

Current Study

In this study, we examined the development of hindsight bias over time and sought to identify individual differences that would predict an initial display of hindsight bias, as well as the "growth," in other words the increase or decrease over time, of the hindsight bias. We examined gender, familiarity, and familiarity squared as predictors of hindsight bias. We also investigated the relationship between need for cognition, self-presentational concerns, and impulsivity and carelessness in social problem solving with hindsight bias.

Our primary hypothesis for the current study was that hindsight bias would be consistently displayed by participants. In addition, in line with the findings of other studies, we hypothesized that hindsight bias would be positively correlated with self-presentational concern, impulsivity/carelessness, and familiarity. We also hypothesized that hindsight bias would be negatively correlated with need for cognition.

Finally, we set out to compare the hypothetical and memory models of hindsight bias. We hypothesized that both models would differ in the magnitude of hindsight bias displayed.

We also hypothesized that individual predictors would differ between models in their ability to explain any variance that existed.

Methods

Participants

Students from the University of Texas at El Paso were recruited from Introduction to Psychology courses for participation in the study. Students selected from 3 forms of the study: a memory study that lasted for 5 weeks, a hypothetical foresight condition lasting for 4 weeks, or a hypothetical hindsight condition that also lasted for 4 weeks. Participants received class credit to reflect their participation in the study.

Fifty-two students joined the hypothetical version of the study (35 in the foresight condition, and 17 in the hindsight condition). Since the sample sizes were small, it would not have been useful to continue with the statistical analysis of the hypothetical condition, and the results from that study are not discussed.

Eighty-one students joined the memory version of the study. Although the memory condition was underpowered, we determined that there were enough participants to make statistical analysis beneficial. The 81 students were composed of two cohorts. The first cohort consisted of 62 students and was administered their first survey between October 22 and 26, prior to the commencement of Week 8 games. The second cohort (composed of 19 students) were administered their initial surveys during the following week (Oct 29 – Nov 2) before the Week 9 games began.

Participants were not included in the sample used for statistical analysis if they did not provide at least 2 weeks of measures of hindsight bias or if their responses were unrealistic (i.e. 100% confidence in each prediction.) The final sample for the study consisted of 65 participants. Of the final sample, the first cohort was comprised of 33 women and 19 men. The mean age was

18.52 (SD = 5.29). The second cohort was made up of 5 women and 8 men with a mean age of 21.38 (SD = 3.80)

Measures

For the initial meeting, participants were administered surveys in person at a campus computer lab via surveymonkey.com. They provided informed consent for their participation. Participants completed a survey of demographics which collected information about their age, sex, education level, ethnicity, and other basic items (see Appendix A).

Additionally, each participant's familiarity with the NFL was determined through a survey in which they were challenged to name the starting quarterback of each of the NFL teams (see Appendix B). The number of quarterbacks correctly named provided an index for familiarity. A number of individual difference measures were also administered and are discussed below.

Self presentational concern. The self-presentational concerns of participants were assessed using the Social Desirability Scale-17 (SDS-17; see Appendix C; Stöber, 1999). The SDS-17 has shown adequate reliability with $\alpha = .74$ (Stöber, 2001). The SDS-17 was developed as an alternative to the traditional Marlowe-Crown scale, which was introduced in 1960 and has become somewhat dated (Stöber, 2001). The SDS-17 correlates significantly with the Marlowe-Crowne scale, $r = .68$, $p < .001$ (Stöber, 2001). However, when researchers asked a sample of students to rate the items from both surveys for social desirability, they found that the ratings of the SDS-17 were substantially higher than those of the Marlowe-Crowne scale with an effect size of $d = .64$ (Stöber, 2001). Since the sample for the present study was composed of college students, the more contemporary design of the SDS-17 was used for this study.

The SDS-17 is comprised of 17 true or false statements regarding participants' behavior in social situations (Stöber, 1999). It is scored by tallying the number of socially desirable responses. Seven of the seventeen statements are reverse keyed (Stöber, 1999). The measure primarily represents the impression management component of social desirability (Musch and Wagner, 2007).

Need for cognition. Participants' need for cognition was assessed by the Need for Cognition Scale, Short Form (NCS; see Appendix D; Cacioppo & Petty, 1984). The scale is made up of 18 items, 9 of which are reverse keyed. Participants are asked to rate the level to which each statement describes them using a Likert scale. Research has shown that the NCS has desirable psychometric properties among college students at UTEP (Culhane, Morera & Hosch, 2004).

Impulsivity. The Impulsivity subscale of Social Problem Solving Inventory- Revised was also administered (SPSI-R; see Appendix E; D'Zurilla, Nezu & Maydeu-Olivares, 1999). The subscale is composed of 10 Likert scale questions. Prior research with the SPSI-R has demonstrated good convergent validity and reliability (D'Zurilla et al., 1999).

Predictions and recall. Finally, participants were asked to make NFL predictions regarding which teams they expected to win each game (see Appendix F). They were shown the image of a helmet and the name of each team and then asked to assign a percentage likelihood of each team winning, adding up to 100%. For example, if the Denver Broncos were playing the New England Patriots and a participant expected the latter to win, they might assign an 80 to the Patriots and a 20 to the Broncos. If they were unsure, they would assign a 50 to each team.

Procedure

Two cohorts of participants participated in the study for a 5-week period. The first cohort consisting of 62 participants began during Week 7 of the 2012-13 NFL season, and the second cohort of 19 students began during Week 8. The first meeting was conducted in person in a computer lab on campus. First, all participants were provided with an informed consent. Next they were guided through a demographics survey, a measure of expertise, as well as measures assessing self-presentational concerns, need for cognition, and impulsivity/carelessness in social problem solving. Measures were counterbalanced each day by administering them in varying sequences to mitigate carryover effects.

Participants were asked to provide an email address where each subsequent week they received an email with a unique link to their survey. First, they were asked to recall the previous week's predictions after being primed with the actual game results. Next, they were asked to provide predictions for the following week's games. Cohort 1's participation lasted from Week 8 through Week 11 of the NFL season and cohort 2's lasted from Week 9 through Week 12.

Measuring the Hindsight Bias

Confidence as an index of hindsight bias. The amount of hindsight bias a participant can display is determined by the accuracy of each initial prediction. The more accurate an initial prediction is, the less hindsight bias a participant can commit. For example, if the Denver Broncos and the Jacksonville Jaguars played a game and a participant assigned the Broncos a 90% chance of winning and the Jaguars a 10% chance of winning and the Broncos won, the participant would only be able to commit 10 points of hindsight bias. If another participant was unsure of who would win and assigned both teams a 50% chance of winning, they would then have the opportunity to commit 50 points of hindsight bias. Thus a participant with more accurate predictions would have the opportunity to commit less hindsight bias than a participant

who was not so accurate. Our measure of hindsight bias takes this into account and is measured as a proportion of total bias committed over the bias possible. Our measure of hindsight bias is calculated as follows:

$$Hindsight\ Bias = \frac{Winning\ Team\ Hindsight - Winning\ Team\ Prediction}{100 - Winning\ Team\ Prediction} \quad (1)$$

If a participant's hindsight recall was less accurate than their initial prediction, the result was a negative hindsight score.

Flips. Additionally, for each game we determined which team the participant favored to win (teams assigned a confidence rating of greater than 50) in both the hindsight and foresight conditions. A "flip" resulted when the participant's foresight prediction favored a team that did not win the game and their hindsight recall is adjusted to favor the actual winning team. For instance, in the previous example, the participant who assigned the Broncos a 90% chance and the Jaguars a 10% chance of winning was expressing a prediction that the Broncos would win. If the Jaguars won and the participant's hindsight recollection was that the Broncos would have a 40% chance and the Jaguars would have a 60% chance of winning, they would have switched their recall to favor the Jaguars and would have committed a flip. In order to create a measure of average flips per week, we divided the number of flips committed by the number of flips possible (the number of incorrect predictions). The result was a proportion between 0 and 1.

Approach to Analyses

Data transformation. Since both resulting measures of hindsight bias were bounded proportions, a logit transformation was performed to stabilize the variance. Our measure of confidence hindsight bias was bounded between -1 and 1 and our flips measure was bounded between the values of 0 and 1. To ensure that the dependent variables were able to be transformed, they were first adjusted to fall between the values of 0 and 1. In the case of the

confidence measure, 1 was added to each value and then the resulting measure was divided by 2. The resulting DV included the values of 0 and 1 which are undefined in logit transformations, so they were multiplied by the constant .99999 so that all values could be computed. Average flips, which already fell between 0 and 1, were simply prepared by being multiplied by the constant .99999.

In order to aid in interpretation of the findings, back transformations of each parameter estimate were taken following the analysis. These values are reported, as well as the logit transformed values.

Maximum Likelihood Estimation. Our model was estimated using Full Information Maximum Likelihood (FIML). Maximum likelihood (ML) estimations use an iterative process to produce parameters that are the most likely to produce the data being analyzed by comparing different estimates using a log likelihood function. As the estimates of the parameters become more likely, the log likelihood function is minimized. When a new iteration of the estimated parameters does not reduce the log likelihood by more than a user-defined convergence criteria (in our study, .001), the iterative process is terminated and the previous set of parameter estimates is reported. In addition to parameter estimates, ML provides estimates of standard errors and fit indices by which the accuracy of the parameter estimates can be evaluated.

Two forms of ML are commonly used: Full Information Maximum Likelihood (FIML) and Restricted Maximum Likelihood (REML). In FIML, the fixed effect parameters are treated as known (resulting in these fixed sample parameters being treated as population parameters), affecting the estimation of random effects. Conversely, REML treats fixed effects as unknown and specifically searches for values that maximize the likelihood of the residuals. FIML, unlike REML, allows the researcher to quantitatively compare nested models that are composed of

fixed and random effects. Given the nature of our study, FIML was chosen as our method of estimation[1].

Growth curve modeling. We modeled change over time using growth curve modeling. In growth curve modeling, we are deriving a regression equation for each individual as well as parameters that describe the entire sample. For example, Equation 2 estimates each Cohort 1 individual's linear growth over time. β_{0j} represents the participant's amount of hindsight bias at Week 8 and β_{1j} represents the individual's change in hindsight bias over time. This model is referred to as the "Level 1" model, as repeated observations are nested within participants

$$Y_{ij} = \beta_{0j} + \beta_{1j}(time_{ij} - \text{Week 8}) + \varepsilon_{ij} \quad (2)$$

The Level 2 portion of this model examines how an individual's intercept and slope differ from the average intercept (denoted γ_{00}) and the average slope (denoted γ_{10}).

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (3)$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

Our examination of flips was straightforward, with four average flip values (one for each week). However, in our assessment of hindsight confidence, we modeled the data using a multiphase multilevel approach which was described by Cudeck and Klebe (2002). This method retains all of the benefits of a multilevel modeling approach by estimating individual parameters for each participant as well as aggregated parameters for each "cluster". The multiphase approach also allows us to also nest time points within "phases" for a more detailed and tailored analysis of our data.

Figure 1 demonstrates the advantages of the multiphase approach to modeling data. If averages of the data for each week (shown in the illustration as an 'x') are used in a multilevel

[1] Maximum Likelihood Estimations were conducted using the MULTILEV module in PRELIS 2.8 (Jöreskog & Sörbom, 2003)

model, the resulting model shows a negative linear trajectory. However, the when the same data are analyzed using a multiphase approach, it becomes apparent that although the means descend in a linear fashion over time, within each week hindsight bias exhibited increases with each successive game.

Using the multiphase approach, the Level 1 model becomes:

$$Y_{ij} = \begin{cases} \alpha_0 + \alpha_1 x_j + \omega_{ij} & x_j < \tau_1 \\ \sigma_0 + \sigma_1 x_j + \Gamma_{ij} & \tau_1 \leq x_j < \tau_2 \\ \Omega_0 + \Omega_1 x_j + \Theta_{ij} & \tau_2 \leq x_j < \tau_3 \\ \delta_0 + \delta_1 x_j + \varphi_{ij} & \tau_3 \leq x_j \end{cases} \quad (4)$$

Our original estimate of ß in Equation 2 is now vector ß = (α_0, α_1, σ_0, σ_1, Ω_0, Ω_1, δ_0, δ_1). Each pair of parameters estimates a separate intercept and slope for any phase, τ_x. This results in a separate linear regression line for each individual, for each phase. In addition, a separate error term is also estimated for each phase in vector $\varepsilon_{ij} = (\omega_{ij}, \Gamma_{ij}, \Theta_{ij}, \varphi_{ij})$. Rather than two fixed effects as in Equation 2, we are now estimating 8 fixed effects and variances for each effect. Analyzing the data in this way, we are able to compare each phase within individuals and assess whether predictors behave differently at separate time points.

Similar to Equation 3, in Level 2 of the model each parameter of vector ß is further defined by its mean and an error term that reflects how each individual differs from the mean. In our model, we allowed for each parameter of vector ß to vary. We also allowed each intercept parameter to covary with its corresponding slope.

Model assumptions. The multilevel model assumes homoscedasticity of the variance, and that error terms at every level are independent and normally distributed. Additionally, the model makes the assumption of linear relationship among the variables.

In order to test the assumption of normality and linearity, we examined Q-Q plots of residuals at each level of the model. If normality held, we assessed whether the variance was

homogenous using Bartlett's test. However, Bartlett's test is sensitive to deviations from normality in data (Levene, 1960). If the normality assumptions did not hold, we used a Levene's test.

Model building approach. For both the confidence as an index of hindsight bias and the flips models, we first assessed an unconditional linear growth model of the data for each measure of hindsight bias in order to detect variance in the intercepts and slopes of each participant. These models included time as a level 1 predictor, but no level 2 predictors. The resulting null model for our flips measure estimated 4 parameters: the intercept, slope, and level 1 and level 2 variances. The confidence null model estimated 40 total parameters: intercepts for each of the 5 weeks of the study, slopes for each of the 5 weeks, variance of the intercept and slope parameters, and covariance between each intercept and its corresponding slope.

Assuming a null model, error terms (u_{0j} and u_{1j} for flips and vector ε_{ij} for confidence as an index of hindsight bias) would be equal to zero. However, a result indicating significant variance in the intercepts in the error terms suggested that there may be individual (level 2) differences to account for the variance. Individual differences were only entered as predictors for the phases where significant variability was present. We did not transform any of the predictor scores, instead retaining the raw scores calculated as directed by the developers of each measure. Additionally, we assessed whether the cohort a participant belonged to explained any existing variance. First, we entered the participant's cohort into the model. Next we entered familiarity, familiarity-squared and sex into level 2 of the model, given previous research that gives us a theoretical basis for their inclusion (Morera et al, 2013). If variability in the intercepts and slopes was still present, we looked at our impulsivity/ carelessness measure, as this variable was found to be a predictor of the hindsight bias in an earlier study (Morera et al, 2013). Finally

we assessed whether the inclusion of need for cognition and self-presentational concerns improved the model and explained any additional unexplained variability in the intercepts and slopes. We also examined interactions, namely expertise and self-presentational concerns.

Results

Model Assumptions

The data was found to violate the assumptions of normality and linearity due to the logit transformation that was taken. As a result, homogeneity of variance was assessed using the Levene's test. For flips, the Levene's test concluded that the assumption homogeneity of variance was upheld, $F(69, 280) = .590$, *ns*. However, for confidence hindsight bias, Levene's test was significant for weeks all weeks (see Table 1).

Individual Difference Measures

The results of the individual difference measures are summarized in Table 2. A frequency distribution of the familiarity scores of the final sample is found in Figure 2. Need for Cognition and Impulsivity and Carelessness scores were moderately, negatively correlated, $r = -.39$, $p < .001$. In addition, Need for Cognition and SDS-17 scores were moderately and positively related, $r = .31$, $p < .05$.

Flips

A multilevel model measuring flips as the outcome variable and the week as a Level 1 predictor is summarized in Table 3. The table first lists the transformed parameter estimate, and then a back transformed estimate which can be interpreted as the proportion of the hindsight bias possible that the average participant actually committed. Analysis revealed that participants did not flip a significant proportion of their incorrect predictions for Week 8, ß = -.95, SE = .68, ns. However, each week, they committed 12% more hindsight flips on average, ß = -1.99, .34, p < .001. In addition, there was no significant Level 2 variance to explain in the proportion of flips committed, $\tau = 0.66$, $SE = 2.31$, *ns*. There was, however, significant Level 1 variance, $\tau = 39.97$, SE = 4.23, p < .001. Since no significant Level 2 variance remained to explain, our investigation

of flips as an index of hindsight bias did not progress beyond the null model. A graphical representation of the average flips per week can be found in Figure 3.

Confidence as an Index of Hindsight Bias

Fixed effects. The fixed effects estimates of the model with confidence as the outcome variable are shown in Table 4.1. As in the results for the flips, all fixed effect transformed parameter estimates and back-transformed estimates are both shown to aid in comprehension. The assessment of confidence as an index of hindsight bias revealed that significant bias was being committed on Game 1 of Week 8 (β = .50, *SE* = .25, $p < .05$; 25% of bias possible), Week 9 (β = .889, *SE* = .39, $p < .001$; 60% of bias possible), and Week 11, β = .34, *SE* = .17, $p < .05$, 13% of bias possible. Additionally, the amount of bias committed on Week 9 reduced by 4% for each progressive game, ß = - 0.08, *SE* = .04, $p < .05$. A graphical representation of the relationships between intercepts and slopes can be found in Figure 4.

Random effects. There was significant level 2 variance in the Week 9 hindsight bias intercept (τ = .99, *SE* = .93, $p < .001$) as well as the Week 9 slope, τ = .07, *SE* = .03, $p < .001$. None of our individual predictors or the participants' cohort were able to account for a significant amount of the variance observed, and our null model became our final model.

We also allowed each intercept to covary with the corresponding slope for its week. On Week 8 (ß = .18, *SE* = .06, $p < .01$), Week 10 (ß = .19, *SE* = .04, $p < .001$), and Week 11 (ß = .05, *SE* = .03, $p < .05$) these covariances were positive and significant, indicating that participants with higher intercepts tended to have steeper slopes.

Level 1 variance was present for all 5 weeks, indicating that even after games were added into the model, individuals still differed from their predicted hindsight bias trajectories. Level 1and 2 variance is shown in detail in Table 4.2.

Discussion

Our hypothesis that hindsight bias would be displayed by participants was confirmed both in terms of flips and confidence as an index of hindsight bias with the exception of Week 7. However, given our small sample size, we were unable to test the hypothesis that the hypothetical and memory models would differ both in the amount of hindsight bias displayed and the ability of individual difference variables to explain variance. In addition, our hypotheses related to the relationship of need for cognition, self-presentational concerns, and familiarity with hindsight bias were not supported contrary to previous research (Blank et al, 2007; Campbell & Tesser, 1983; Gray et al, 2007; Marks Knoll, 2009; Morera et al, 2013; Musch and Wagner, 2007; Verplanken & Pieters, 1988).

The lack of support of our hypotheses is most likely a direct result of our lack of power. To achieve a power of .8 for the memory condition, we needed a sample of no less than 98 subjects. Our sample size of 65 did not meet this requirement, and as a result, we only achieved a power of about .55 (Raudenbush et al, 2011). With inadequate power, we may have been unable to detect an effect that was actually present. The lack of power also has serious implications for the generalizability of our findings. However, the study did yield some interesting results that are worthy of further investigation.

Perhaps the most notable contribution of the current study is the use of multiphase mixed effects models to analyze longitudinal memory model (Cudeck & Klebe, 2002). Using this approach, we were able to model the data in detail that, to our knowledge, has not yet been achieved in the literature. In the future, similar modeling could be used to investigate complex and delicate relationships between the display of hindsight bias and individual predictors. The

multiphase model could also be extended to the flips outcome variable, resulting in a logistic regression, although in the current study we analyzed the weekly average flips.

In terms of our findings, weekly trajectories of confidence as an index of hindsight bias for each progressive game are interesting to note. During Week 8, participants showed significant initial bias and increasing bias for each successive game that they estimated. However, in the remaining weeks with significant bias, the amount of bias displayed decreased with every game with the exception of Week 10. This decrease in bias was unexpected, as previous research has shown that bias increases with each successive game (Morera et al, 2013).

Several reasons may exist that could explain the disparity in our results. The first is that the relationship could, in fact, be non-linear. The games were presented for prediction and recall in the same order. Primacy and recency effects could have resulted in the earliest and latest predictions being more accurately recalled (resulting in less bias) than predictions in the middle of the survey (Murdoch, 1962).

In our study, Week 8 and 9 predictions included Thursday night games. However, as the study progressed, we decided not to analyze Thursday night games due to low participation. In this way, we could run participants on Fridays after the Thursday night games had already been concluded. As a result, the surveys for Weeks 8 and 9 included Thursday night games, but they were eventually removed from the analysis. This means that the intercept for those weeks was actually the hindsight bias exhibited on the second game and primacy effects may not have been displayed in the model. This could have resulted in a much higher intercept than would have been present had the Thursday night games been preserved in the study. This could make the results in Week 9 appear to decline in a linear fashion, when actually they more closely resemble an inverse parabola with early and late games being met with less hindsight bias than those in the

middle. In the future, it would be worthwhile to analyze repeated measure hindsight bias data taking the primacy and recency effects into account and modeling a non-linear relationship.

Additionally, for the memory model of this study, there was no control group to assess whether accuracy in memory was mitigating the effects of hindsight bias. It is possible that the negligible findings are the result of high accuracy in participants' recalls. In future studies, it would be appropriate to include a second memory condition where participants are not shown the outcome of the games and are asked to reproduce their predictions by memory alone.

Moreover, there may be specific games that are more likely to be accurately recalled. For example, participants might pay special attention to games that receive more attention from the media or games or feature the individual's favorite team, resulting in less hindsight bias and a more accurate recollection of predictions. One benefit of the multiphase mixed effects approach is that in the future, we could include the popularity of teams or individual games in the model (Cudeck & Klebe, 2002). The multiphase model could also be used to model which games participants care about as a level 1 random effect.

The addition of a control condition could also be used to control for practice effects, where the participant becomes more familiar with the task as time passes and is able to more efficiently perform with each trial (Petersen, Van Mier, Fiez, & Raichle, 1998). Perhaps as they learn that every prediction is followed by a recall, they are developing more effective techniques for recalling the information.

The current study only took time-invarying covariates into account, taking measures of individual differences only on the first week. It is possible that these effects might vary from week to week and would be able to better explain fluctuations in the hindsight bias over time.

Need for cognition has been shown to be stable over an individuals lifetime and to maintain reliability between tests and retests (Cacioppo & Petty, 1982; Sadowski & Gulgoz, 1992). However, one meta analysis suggested that, particularly in computer administered surveys, the effects of impression management on response patterns reduced over time (Dwight & Feigelson, 2013). It is possible that by the fourth week, self-presentational concerns did not play as large of a role in responses, resulting in lower levels of hindsight bias.

Additionally, as individuals are making predictions of NFL games, perhaps they are paying more attention to game outcomes. This provides reason to believe that with each subsequent week, participants' familiarity with NFL teams and game outcomes increased. However, the rate of increase would differ among participants depending on how much they engaged with the NFL outside of taking surveys. It would be worthwhile to study whether the change in familiarity over time can predict changes in hindsight bias.

The more familiar with the NFL that a participant becomes, the more likely that they are able to make predictions based on prior knowledge and reason. Studies have shown that asking participants to proactively engage in reasoning for why an experimental outcome has occurred results in reduced magnitude and frequency of hindsight bias displayed (Slovic & Fischoff, 1977). It is possible that the increased familiarity with the NFL joined with the knowledge that the participant will be asked to recall their prediction prompts them to engage in more in depth reasoning for their prediction.

Another factor that was not taken into account into the current study was the occurrence of unlikely outcomes. A popular cognitive process model, the Causal Model Theory (CMT), states that when participants are told the outcome of an event, they are naturally drawn to construct an explanation of why the outcome occurred by searching for causal antecedents

(Blank & Nestler, 2007). According to the CMT model, the hindsight bias is a result of a causal connection being established (Blank & Nestler, 2007). However, if the causal connection is not able to be made, hindsight bias does not occur. Researchers have found that the metacognitive experience of surprise (an outcome for which no causal antecedent is found) diminishes the display of hindsight bias (Sanna and Schwarz, 2006; Yopchick & Kim, 2012; Blank & Nestler, 2007). In games where the unfavored team won the game, participants may have displayed less bias. Future studies taking the likelihood of the outcome into account might reveal interesting effects.

Finally, how does the event in question affect the level of hindsight bais being displayed? There are obviously major differences between hindsight bias related to a football game versus a natural disaster such as Hurricane Katrina. The meta-analysis conducted by Guilbault et al (2004) assessed whether the event or task type, positive or negative valence of the outcome, or personal relevance of the event could significantly predict hindsight bias displayed. They found that personal relevance was not a significant predictor of hindsight bias, $Q(1) = 1.88$, *ns* (Guilbault et al, 2004). However, the valence of the outcome ($Q(2) = 21.11$, $p < .001$) and the type of event or task ($Q(2) = 24.75$, $p < .001$) were significant predictors, with neutral almanac questions yielding higher levels of hindsight bias than positive or negative outcomes of real world events or case histories (Guilbault et al, 2004). Apart from the aforementioned study regarding the Chernobyl disaster, there has been very little in the literature addressing hindsight bias in events of great magnitude (Verplanken & Pieters, 1988).

Also, Hurricane Katrina arguably marks a turning point in United States history in regards to hurricane preparedness and public attentiveness. Since 2005, a storm has scarcely touched the coast of Louisiana didn't conjure comparisons to Katrina in the media. Hurricane

Katrina has effectively become the standard by which many Americans measure the severity of each successive hurricane. Does the unique nature of Hurricane Katrina as a game-changer affect the hindsight bias exhibited after the storm and the lack thereof following hurricanes since? This is one of the many worthwhile research questions to address regarding the impact of the nature of the event on the bias exhibited by individuals.

Despite the difficulties encountered during the data collection for the current study, it provides a beginning to what is an important field of research. As New Orleans witnessed in 2005, the effects of hindsight bias can be devastating. Although in our study, the outcomes were not nearly as extreme as the outcomes in a natural disaster, by studying hindsight bias in a controlled context, we are able to begin distilling underlying predictors and trends of its occurrence. As our understanding is broadened through these studies, they will inform our approach to studying hindsight bias on larger scales.

References

Arkes, H. R., Faust, D., Guilmette, T. J., & Hart, K. (1988). Eliminating the Hindsight Bias. *Journal of Applied Psychology*, *73*(2), 305-307. doi:0021-9010/88

Arkes, H. R., Wortmann, R. L., Saville, P. D., & Harkness, A. R. (1981). Hindsight bias among physicians weighing the likelihood of diagnoses. *Journal of Applied Psychology, 66*(2), 252-254. doi:10.1037/0021-9010.66.2.252

Ash, I. K. (2009). Surprise, Memory, and Retrospective Judgment Making: Testing Cognitive Reconstruction Theories of the Hindsight Bias Effect. *Journal Of Experimental Psychology. Learning, Memory & Cognition*,*35*(4), 916-933. doi:10.1037a0015504

Blake, E. S., Jarrell, J. D., Mayfield, M., Rappaport, E. N., & Landsea, C. W. (2005). US mainland hurricane strikes by state, 1851-2004.

Blank, H., & Nestler, S. (2007). Cognitive process models of hindsight bias. *Social Cognition, 25*(1), 132-146, doi:10.1521/soco.2007.25.1.132

Blank, H., Musch, J., & Pohl, R. F. (2007). Hindsight bias: On being wise after the event. *Social Cognition*, 25(1), 1-9, doi:10.1521/soco.2007.25.1.132

Bonds-Raacke, J. M., Fryer, L. S., Nicks, S. D., & Durr, R. T. (2001). Hindsight bias demonstrated in the prediction of a sporting event. *The Journal Of Social Psychology*, *141*(3), 349-352. doi:10.1080/00224540109600556

Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306-307.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197.

Campbell, J. D., & Tesser, A. (1983). Motivational interpretations of hindsight bias: An individual difference analysis. *Journal Of Personality*, *51*(4), 605. doi:10.1111/1467-6494.ep7380582

Christensen-Szalanski, J., & Willham, C. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes, 48,* 175-168. doi:10.1016/0749-5978(91)90010-Q

Cudeck, R., & Klebe, K. J. (2002). Multiphase mixed-effects models for repeated measures data. *Psychological Methods, 7*(1), 41.

Culhane, S. E., Morera, O. F., & Hosch, H. M. (2004). The factor structure of the need for cognition short form in a Hispanic sample. *The Journal of Psychology*, *138*(1), 77-90.

Dwight, S. A., & Feigelson, M. E. (2000). A quantitative review of the effect of computerized testing on the measurement of social desirability. *Educational and Psychological Measurement, 60*(3), 340-360. doi:10.1177/00131640021970583

Experimental Psychology: Human Perception and Performance, 3, 544-551. doi:10.1037//0096-1523.3.4.544

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. Journal Of Experimental Psychology: Human Perception And Performance, 1(3), 288-299. doi:10.1037/0096-1523.1.3.288

Fischhoff, B., Beyth, R. (1975). I knew it would happen: Remembered probabilities of once—future things. *Organizational Behavior and Human Performance,13*(1), 1-16, doi:10.1016/0030-5073(75)90002-1.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition

Grace, S. (2012, September 2). Government does its job during Hurricane Isaac: Stephanie Grace. Retrieved from Times-Picayune: http://www.nola.com/opinions/index.ssf/2012/09/government_does_its_job_during.html

Gray, R., Beilock, S. L., & Carr, T. H. (2007). "As soon as the bat met the ball, I knew it was gone": Outcome prediction, hindsight bias, and the representation and control of action in expert and novice baseball players. *Psychonomic Bulletin & Review, 14*(4), 669-675. doi:10.3758/BF03196819

Guilbault, R. L., Bryant, F. B., Brockway, J., & Posavac, E. J. (2004). A Meta-Analysis of Research on Hindsight Bias. *Basic & Applied Social Psychology*, *26*(2/3), 103-117. doi:10.1207/s15324834basp2602&3_1

Harley, E. M. (2007). HINDSIGHT BIAS IN LEGAL DECISION MAKING. *Social Cognition, 25*(1), 48-63.

Harley, E. M., Carlsen, K. A., & Loftus, G. R. (2004). The 'Saw-It-All-Along' Effect: Demonstrations of Visual Hindsight Bias. *Journal Of Experimental Psychology: Learning, Memory, And Cognition, 30*(5), 960-968. doi:10.1037/0278-7393.30.5.960

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin,107*(3), 311-327. doi:10.1037/0033-2909.107.3.311

Leary, M. R. (1981). The distorted nature of hindsight. *The Journal Of Social Psychology, 115*(1), 25-29.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model.*Psychological Bulletin*, *107*(1), 34-47. doi:10.1037/0033-2909.107.1.34

Levene, H. (1960). Robust tests for equality of variances1. *Contrib Prob Stat: Essays Honor Harold Hotel, 2*, 278.

Louie, T. A. (1999). Decision makers' hindsight bias after receiving favorable and unfavorable feedback. *Journal Of Applied Psychology*, *84*(1), 29-41. doi:10.1037/0021-9010.84.1.29

Louie, T. A., Curren, M. T., & Harich, K. R. (2000). 'I knew we would win': Hindsight bias for favorable and unfavorable team decision outcomes. *Journal Of Applied Psychology*, *85*(2), 264-272. doi:10.1037/0021-9010.85.2.264

Louie, T. A., Rajan, M. N., & Sibley, R. E. (2007). Tackling the Monday-morning quarterback: Applications of hindsight bias in decision-making settings. *Social Cognition*, *25*(1), 32-47. doi:10.1521/soco.2007.25.1.32

Mark, M., Reiter Boburka, R., Eyssell, K., Cohen, L., & Mellor, S. (2003). "I couldn't have seen it coming": The impact of negative self-relevant outcomes on retrospections about foreseeability. *Memory*, *11*(4/5), 443.

Marks Knoll, M. A. (2009). The Effects of Expertise on Hindsight Bias.

Martin, K. A., & Leary, M. R. (2001). Self-presentational determinants of health risk behavior among college freshmen. *Psychology & Health*, *16*(1), 17-27. doi:10.1080/08870440108405487

Morera, O. F., Stokes, S. M., Duran, M. (2013, May). *Hindsight Bias: A Longitudinal Study.* Presentation at annual conference for the Oklahoma Judgment and Decision Making Society, Oklahoma City, OK.

Murdoch Jr, B. B. (1962). The serial position effect of free recall1. *Journal of Experimental Psychology*, *64*(5), 482-488.

Musch, J. (2003). Personality differences in hindsight bias. Memory, 11(4/5), 473.

Musch, J., & Wagner, T. (2007). Did Everybody Know it All Along? A Review of Individual Differences In Hindsight Bias. *Social Cognition*, *25*(1), 64-82. doi:10.1521/soco.2007.25.1.64

Nestler, S., & Egloff, B. (2009). Increased or Reversed? The Effect of Surprise on Hindsight Bias Depends on the Hindsight Component. *Journal Of Experimental Psychology. Learning, Memory & Cognition*, *35*(6), 1539-1544. doi:10.1037/a007006

Nestler, S., & von Collani, G. (2008). Hindsight bias, conjunctive explanations and causal attribution. *Social Cognition*, *26*(4), 482-493

Osberg, T. M. (1987). The Convergent and Discriminant Validity of the Need for Cognition Scale. *Journal Of Personality Assessment*,*51*(3), 441.

Pennington, D. C. (1981). Being wise after the event: An investigation of hindsight bias. *Current Psychological Research*, *1*(4), 271-282. doi:10.1007/BF03186737

Pezzo, M. V., & Beckstead, J. W. (2008). The effects of disappointment on hindsight bias for real-world outcomes. *Applied Cognitive Psychology*, *22*(4), 491-506. doi:10.1002/acp.137

Petersen, S. E., Van Mier, H., Fiez, J. A., & Raichle, M. E. (1998). The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences, 95*(3), 853-860.

Raudenbush, S. W., et al. (2011). Optimal Design Software for Multi-level and Longitudinal Research (Version 3.01) [Software]. Available from www.wtgrantfoundation.org.

Sadowski, C. J., & Gulgoz, S. (1992). Internal consistency and test-retest reliability of the need for cognition scale. *Perceptual and Motor Skills, 74*(2), 610-610.

Sanna, L. J., & Schwarz, N. (2006). Metacognitive Experiences and Human Judgment: The Case of Hindsight Bias and Its Debiasing. *Current Directions In Psychological Science (Wiley-Blackwell), 15*(4), 172-176. doi:10.1111/j.1467-8721.2006.00430.x

Shanteau, J. (1988). Psychological characteristics and strategies of expert decision makers. *Acta Psychologica, 68*(1), 203-215.

Shanteau, J. (1992). The psychology of experts. *Expertise and Decision Support, 5*, 11-23.

Shanteau, J., Weiss, D. J., Thomas, R. P., Pounds, J., & Hall, B. (2003). How can you tell if someone is an expert? empirical assessment of expertise. *Emerging Perspectives on Judgement and Decision Research,* , 620-639.

Slovic, P. & Fischoff, B. (1977). On the psychology of experimental surprises. *Journal of*

Spybrook, J., et al. (2011). Optimal Design for Longitudinal and Multilevel Research: Documentation for the Optimal Design Software Version 3.0. Available from www.wtgrantfoundation.org.

Staff, C. W. (2012, Sep 12). At least 13,000 Louisiana homes damaged by Hurricane Isaac. Retrieved from CNN U.S.: www.cnn.com/2012/09/04/us/severe-weather

Stöber, J. (2001). The Social Desirability Scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal Of Psychological Assessment*, *17*(3), 222-232. doi:10.1027//1015-5759.17.3.222

Verplanken, B., & Pieters, R. G. (1988). Individual differences in reverse hindsight bias: I never thought something like Chernobyl would happen. Did I?. *Journal Of Behavioral Decision Making*, *1*(3), 131-147. doi:10.1002/bdm.3960010302

Yopchick, J. E., & Kim, N. S. (2012). Hindsight bias and causal reasoning: a minimalist approach. *Cognitive Process*(13), 63-72. doi:10.1007/s10339-011-0414-z

Table 1 *Levene's test results for confidence as an index of hindsight bias*

Week	F- value	P-value
8	10.04	.000
9	7.68	.000
10	15.92	.000
11	6.45	.000
12	4.165	.000

Note: Degrees of freedom for all weeks: (64, 780)

Table 2 *Individual Difference Measures*

Individual Difference	Range Possible	Mean	SD
Familiarity	0 to 32	10.59	10.02
Need for Cognition	-68 to 68	11.37	16.93
SDS-17	0 to 16	8.63	2.80
Impulsivity and Carelessness	0 to 40	9.21	5.96

Table 3 *Hindsight Bias Flip Model Summary*

Parameter Estimate	Greek Parameter	Raw Estimate (%)	Standard Error	Z-test	P-value
Fixed Effects (ß)					
Intercept	- .95	-	-	-	NS
Weeks	**-1.99**	**.12**	**.34**	**-5.79**	**.000**
Random Effects (τ)					
Level 2- Intercept	.66	-	-	-	NS
Level 1- Intercept	**39.97**	**-**	**4.23**	**9.45**	**.000**

Notes: All bolded measures are significant at the $p < .05$ level

Table 4.1
Confidence as an index of Hindsight Bias: Fixed Effects

Parameter Estimate	Greek Parameter (ß)	Raw Estimate (%)	Standard Error	Z-test	P-value
Fixed Effects					
Intercepts					
Week 5	**.51**	**.25**	**.20**	**2.50**	**.01**
Week 6	**1.39**	**.60**	**.30**	**4.69**	**.000**
Week 7	.09	-	-	-	NS
Week 8	.26	-	-	-	NS
Week 9	.12	-	-	-	NS
Slopes					
Week 5	.03	-	-	-	NS
Week 6	**- .08**	- .04	**.04**	**-2.25**	**.02**
Week 7	.02	-	-	-	-
Week 8	- .01	-	-	-	-
Week 9	- .01	-	-	-	NS

Notes: All bolded measures are significant at the $p < .05$ level

Table 4.2

Confidence as an index of hindsight bias random effects

Parameter Estimate	Greek Parameter (τ)	Standard Error	Z-test	P-value
Random Effects Level 2				
Intercepts				
Week 5	-.99	-	-	NS
Week 6	**3.23**	**1.00**	**3.24**	**.001**
Week 7	-1.05	-	-	NS
Week 8	- .15	-	-	NS
Week 9	- .01	-	-	NS
Slopes				
Week 5	- .02	-	-	NS
Week 6	**.03**	**.01**	**2.30**	**.02**
Week 7	- .01	-	-	NS
Week 8	- .01	-	-	NS
Week 9	.00	-	-	NS
Intercept/Slope Covariance				
Week 5	**.17**	**.06**	**2.93**	**.003**
Week 6	- .22	-	-	NS
Week 7	**.19**	**.04**	**4.31**	**.000**
Week 8	**.05**	**.03**	**1.97**	**.05**
Week 9	.00	-	-	NS
Random Effects Level 1				
Week 5	**7.72**	**.47**	**16.55**	**.000**
Week 6	**8.46**	**.45**	**18.61**	**.000**
Week 7	**5.19**	**.30**	**17.17**	**.000**
Week 8	**3.14**	**.20**	**15.92**	**.000**
Week 9	**.22**	**.03**	**6.71**	**.000**

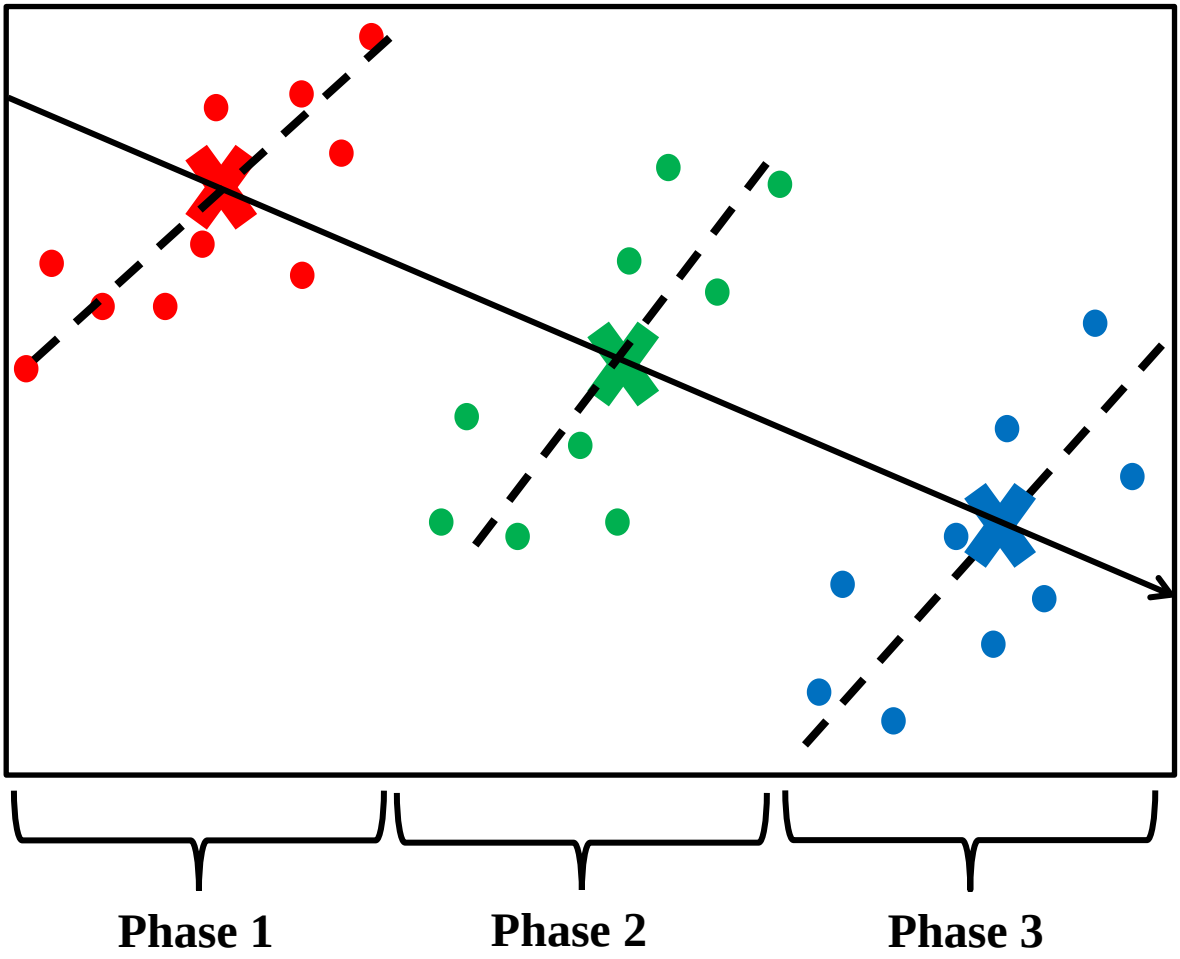


Figure 1. *Hypothetical data illustrating the benefit of multiphase versus aggregated models. Each dot represents a data point and each "x" represents the mean for each phase. The solid line represents the regression line derived from the means of each phase and the dashed lines represent the regression lines of each phase.*

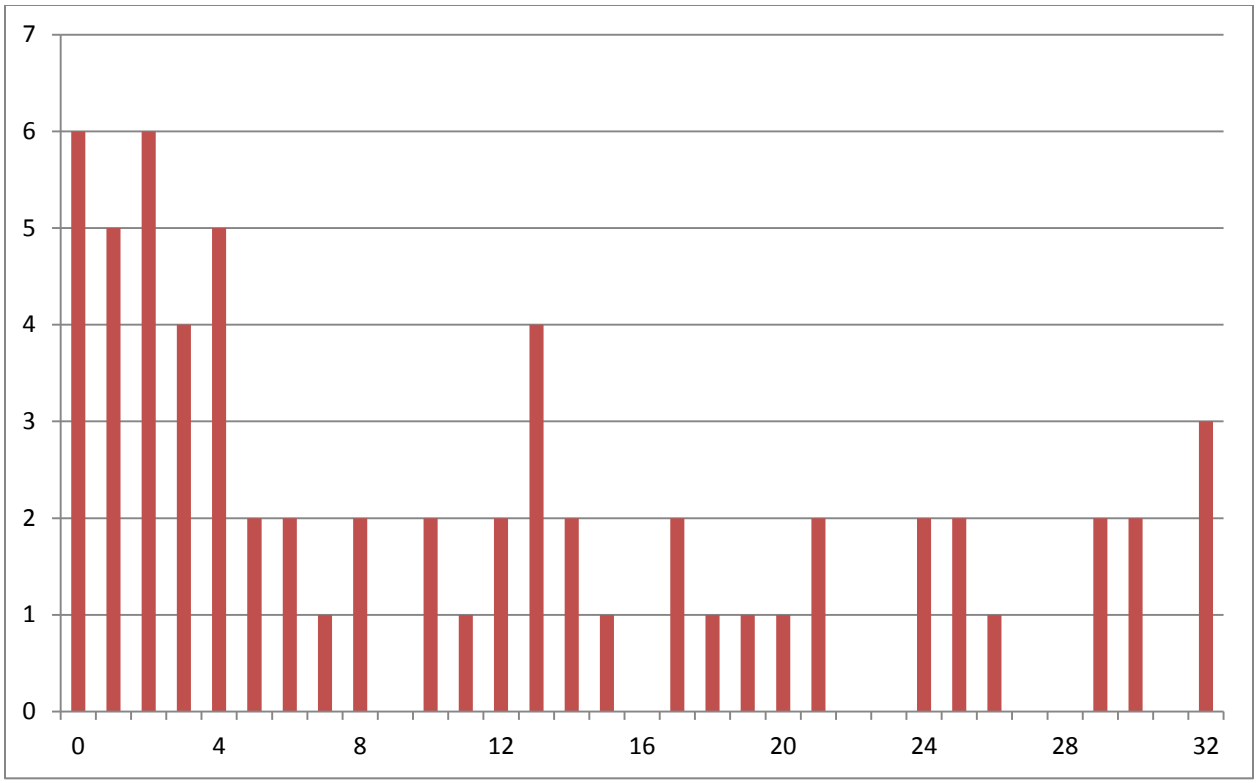


Figure 2. *Frequency of familiarity scores in the final sample.*

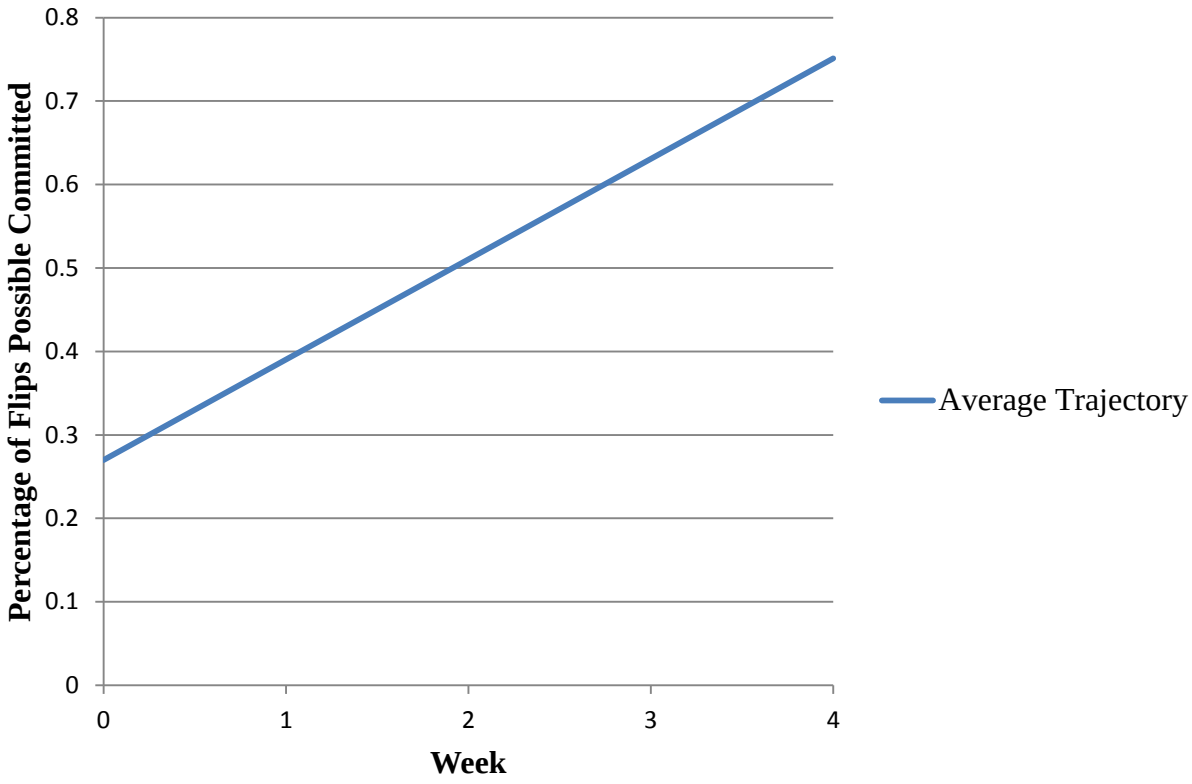


Figure 3. *Average hindsight bias flips committed per week.*

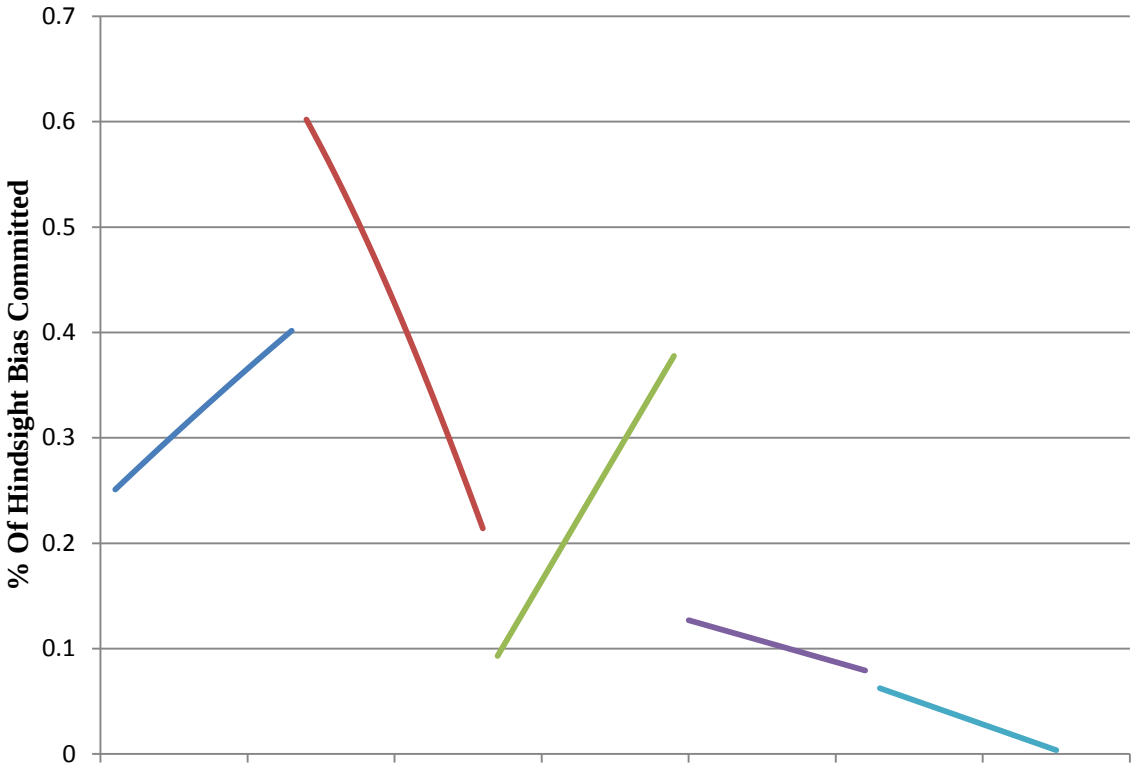


Figure 4. *Average confidence as an index of hindsight bias committed during each week. Each week's trajectory is determined by the growth or decline of bias for each successive prediction/recall beginning with the first game and ending with the final game.*

Appendix A

Demographics

1. Age

2. Gender
__Male
__Female

3. Ethnicity/Race
__Native American
__Pacific Islander
__Black or African American
__Hispanic American
__Mexican National
__Black non-American
__Asian American
__American Indian
__South Asian (India, Pakistan)
__East Asian (Chinese, Japanese, Korean)
__White, Caucasian, European, not Hispanic
__Other (please specify)_____

4. Marital Status
__Single
__Married
__Divorced
__Separated

5. Your personal household income (this includes only your income)
__Between $10,000/year and $19,999/year
__Between $20,000/year and $29,999/year
__Between $40,000/year and $49,999/year
__ Between $50,000/year and $59,999/year
__No income
__Other (please specify)_____

6. Family household income (this includes the income of all household members):
__Between $10,000/year and $19,999/year
__Between $20,000/year and $29,999/year
__Between $30,000/year and $39,999/year
__Between $40,000/year and $49,999/year
__Between $50,000/year and $59,999/year
__Between $60,000/year and $69,999/year
__Between $70,000/year and $79,999/year
__$80,000/year or more
__Other (please specify)

7. Classification
__Freshman
__Sophomore
__Junior
__Senior
__Graduate
__Non-degree graduate
__Staff

8. The academic program you are in:

__Business
__Education
__Engineering
__Health Sciences
__Liberal Arts
__Nursing
__College of Science
__University College
__Undeclared

9. What is your first language?

__English
__Spanish
__Other (please specify)

10. What is your second language?

__English
__Spanish
__Don't speak other language
__Other (please specify)_____

11. In which country have you lived most of your life?

__USA
__Mexico
__Other (please specify) _____

Appendix B

Measure of Familiarity

Below you will find a list of NFL teams along with their logos in no particular order. In the space provided, please identify all of the starting quarterbacks for the NFL teams that you can think of.

Tampa Bay _____

Buffalo Bills _____

Arizona Cardinals _____

Baltimore Ravens _____

Cleveland Browns _____

Denver Broncos _____

Indianapolis Colts _____

Seattle SeaHawks _____

Minnesota Vikings _____

New York Giants _____

St Louis Rams _____

Kansas City Chiefs _____

Washington Redskins

Green Bay Packers

New England Patriots

Cincinnati Bengals

Dallas Cowboys

Jacksonville Jaguars

Carolina Panthers

Atlanta Falcons

New York Jets

Detroit Lions

Miami Dolphins

San Diego Chargers

Philadelphia Eagles

Houston Texans

Oakland Raiders

New Orleans Saints

San Francisco 49er's

Tennessee Titans

Chicago Bears

Pittsburgh Steelers

Appendix C

Social Desirability Scale-17

Below you will find a list of statements. Please read each statement carefully and decide if that statement describes you or not. If it describes you, check the word “true”; if not, check the word “false.”

1. I sometimes litter
2. I always admit my mistakes openly and face the potential negative consequences.
3. In traffic I am always polite and considerate of others.
4. I always accept others' opinions, even when they don't agree with my own.
5. I take out my bad moods on others now and then.
6. There has been an occasion when I took advantage of someone else.
7. In conversations I always listen attentively and let others finish their sentences.
8. I never hesitate to help someone in case of emergency.
9. When I have made a promise, I keep it – no ifs, ands or buts.
10. I occasionally speak badly of others behind their back.
11. I would never live off other people.
12. I always stay friendly and courteous with other people, even when I am stressed out.
13. During arguments I always stay objective and matter-of-fact.
14. There has been at least one occasion when I failed to return an item that I borrowed.
15. I always eat a healthy diet.
16. Sometimes I only help because I expect something in return.

Answer categories are “true” (1) and “false” (0). Items 1, 4, 6, 7, 11, 15, and 17 are reverse keyed.

Appendix D
Need For Cognition Scale- Short Form

1. I would prefer complex to simple problems.

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

2. I like to have the responsibility of handling a situation that requires a lot of thinking.

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

3. Thinking is not my idea of fun.*

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.*

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

5. I try to anticipate and avoid situations where there is likely a chance I will have to think in depth about something.*

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

6. I find satisfaction in deliberating hard and for long hours.

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

7. I only think as hard as I have to.*

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

8. I prefer to think about small, daily projects to long-term ones.*

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

9. I like tasks that require little thought once I've learned them.*

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

10. The idea of relying on thought to make my way to the top appeals to me.

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

11. I really enjoy a task that involves coming up with new solutions to problems.

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

11. Learning new ways to think doesn't excite me very much.*

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

12. I prefer my life to be filled with puzzles that I must solve.

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

13. The notion of thinking abstractly is appealing to me.

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

14. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

15. I feel relief rather than satisfaction after completing a task that required a lot of mental effort.*

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

16. It’s enough for me that something gets the job done; I don’t care how or why it works.*

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

17. I usually end up deliberating about issues even when they do not affect me personally

(-4)-------(-3)--------(-2)--------(-1)--------(0)--------(1)---------(2)---------(3)---------(4)

Very strong agreement	Strong agreement	Moderate agreement	Slight agreement	Neither agreement nor disagreement	Slight disagreement	Moderate disagreement	Strong disagreement	Strong disagreement

Appendix F

NFL Week 6 Predictions

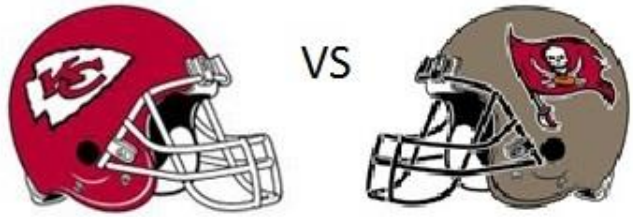
For the following NFL games, please estimate how likely each team is to win by providing a percentage. Your percentages should add up to 100%. For example, if you believe that one team is certain to win, you would assign that team a 100% likelihood and the other team a 0% likelihood. In contrast, if you feel both teams have an equal chance of winning, you would assign them both a 50% likelihood. Your estimates may be any whole number between 0 and 100 (i.e. 73 and 27, 80 and 20, etc).

IRB Reviewers: Please note that the team records are hypothetical. Once the identity of the teams is known, this file will be changed to reflect the records of the teams playing each game.

Game 1: **Pittsburgh Steelers at the Tennessee Titans (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
Pittsburgh Steelers ___
Tennessee Titans ___

Game 2: **Kansas City Chiefs (4-0) at the Tampa Bay Buccaneers (1-3)**



Expressed in percentages, how likely do you estimate the following teams are to win the game?
Kansas City Chiefs ___
Tampa Bay Buccaneers ___

Game 3: **Oakland Raiders at the Atlanta Falcons (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
Oakland Raiders ___
Atlanta Falcons ___

Game 4: **Indianapolis Colts at the New York Jets (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
Indianapolis Colts ___
New York Jets ___

Game 5: **St Louis Rams at the Miami Dolphins (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
St Louis Rams ___
Miami Dolphins ___

Game 6: **Dallas Cowboys at the Baltimore Ravens (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
Dallas Cowboys ___
Baltimore Ravens ___

Game 7: **Detroit Lions at the Philadelphia Eagles (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
Detroit Lions ___
Philadelphia Eagles ___

Game 8: **Cincinnati Bengals (2-2) at the Cleveland Browns (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
Cincinnati Bengals ___
Cleveland Browns ___

Game 9: **New England Patriots (2-2) at the Seattle Seahawks (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
New England Patriots ___
Seattle Seahawks ___

Game 10: **Buffalo Bills (2-2) at the Arizona Cardinals (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
Buffalo Bills ___
Arizona Cardinals ___

Game 11: **Minnesota Vikings (2-2) at the Washington Redskins (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
Minnesota Vikings ___
Washington Redskins ___

Game 12: **New York Giants (2-2) at the San Francisco 49er's (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
New York Giants ___
San Francisco 49er's ___

Game 13: **Green Bay Packers (2-2) at the Houston Texans (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
Cincinnati Bengals ___
Cleveland Browns ___

Game 13: **Denver Broncos (2-2) at the San Diego Chargers (1-3)**

Expressed in percentages, how likely do you estimate the following teams are to win the game?
Denver Broncos ___
San Diego Chargers ___

Curriculum Vita

Sonya M. Stokes (née Carpenter) was born on March 21, 1984. The oldest child of Fred and Sandy Carpenter, she graduated from Western High School in Russiaville Indiana. Sonya received her Bachelor of Arts in Psychology from the University of Texas at El Paso (UTEP) in 2012, graduating Cum Laude. In addition, she was awarded an Outstanding Research Award for her honors thesis titled "Hindsight Bias: A Longitudinal Study." In 2012, she was accepted into the Masters of Experimental Psychology program at UTEP under the supervision of Dr. Osvaldo F. Morera where she worked as a graduate assistant for the Hispanic Health Disparities Research Center.

Sonya will continue her education in the fall of 2013 at the University of Houston where she will pursue a Ph.D. in Industrial/Organizational Psychology. Her research interests include growth curve modeling and multilevel structural equations modeling applied in organizational settings.

Permanent address: 6420 Casper Ridge
El Paso, Texas 79912

This thesis/dissertation was typed by Sonya M. Stokes, B.A.