

Match madness: Probability matching in
the prediction of the NCAA basketball tournament

Sean M. McCrea

University of Konstanz

Edward R. Hirt

Indiana University Bloomington

Address for correspondence:

Sean M. McCrea
Department of Psychology
Fach D-39
University of Konstanz
78457 Konstanz, Germany

Email: Sean.McCrea@uni-konstanz.de

Telephone: +49 7531 88-5103

Fax: +49 7531 88-3286

Abstract

Every year, billions of dollars are spent gambling on the outcomes of the NCAA men's basketball tournament. The current work examines how individuals make predictions for one of the most popular forms of betting, tournament pools, in which individuals must correctly predict as many games in the tournament as possible. We demonstrate that individuals predict more upsets (i.e., wins by a higher-seeded team) than would be considered rational by a normative choice model, and that individuals are no better than chance at doing so. These predictions fit a pattern of probability matching, in which individuals predict upsets at a rate equal to past frequency. This pattern emerges because individuals believe the outcomes of the games are non-random and therefore predictable.

Keywords: Judgment, Sports gambling, Social prediction

The focus of the current work is “March Madness,” as the NCAA men’s college basketball championship tournament has come to be known. Sixty-four of the best teams in the country are selected to play in a six-round, single-elimination tournament to decide the national champion. This tournament has become one of the most popular sporting events of the year, and the focus of a great deal of prognostication. More than \$2.5 billion are spent each year in legal and illegal gambling on this basketball tournament alone (Clarke, 2005). One of the most popular forms of competition (for money or for pride) involves “pools,” in which individuals predict the outcome of all 63 games in the tournament, gaining points based on the number of correct predictions.

There are a number of interesting features of this tournament. First, teams are sorted into four groups (or “regions”) of sixteen teams and ranked (or “seeded”) by the tournament organizers. In the first round, the best, lowest seeded team (1) in a group plays the worst, highest seeded team (16) in the group, the second-lowest seeded team (2) plays the second-highest seeded team (15), and so on. As a result, in each group there are eight first-round games, ranging from an almost sure win by the lowest seeded team to a roughly even match between the eighth and ninth seeded teams. Thus, in predicting the outcome of these games in a typical pool, individuals are asked to predict events with a range of probabilities, but that have equal utility in terms of gaining points and winning.

As the name “March Madness” implies, the attractive as well as the frustrating aspect of predicting the outcomes of these games is the upset, when the lower seeded favored team is beaten by a relative unknown, higher seeded “Cinderella.” To achieve a perfect score, individuals must correctly predict these upsets. Previous studies of the tournament (e.g., Kaplan & Garstka, 2001; Lane & Damiano, 1994) have tended to focus on developing models to better predict the game outcomes rather than examining the ability of individuals to do so. To our knowledge, there have been only a few studies assessing actual prediction

performance, showing for example that individuals perform more poorly when asked to explain their reasoning (Halberstadt & Levine, 1999), and that they insufficiently consider the competitive aspect of tournament pools (Metrick, 1996). The present studies have a different aim, seeking to examine whether individuals attempt to successfully predict upsets, how they might make upset predictions, and whether they can outperform chance in making them. Specifically, we sought to demonstrate that individuals predict upsets in an attempt to achieve a perfect score, but that doing so is extremely difficult, such that even knowledgeable individuals fail to perform better than chance in their upset predictions.

Past tournament results

Figure 1 presents the actual results for each of the eight first-round games (i.e., 1 seed vs. 16 seed, 2 seed vs. 15 seed, etc.) collapsed across the four groups (i.e., regions) for the years 1985-2005. What can be seen is that, with the possible exception of the 8-9 game in which the higher seeded 9 seed has won slightly more often, the tournament seedings are good predictors of the outcome of the games. Indeed, a number of investigations have demonstrated that making predictions on the basis of rankings made by other experts (e.g., polling from sports reporters or coaches) or computer formulas, Las Vegas odds, or more complex statistical modeling all produce about the same level of performance on average as simply predicting the lower seed to win (Kaplan & Garstka, 2001; Lane & Damiano, 1994). Thus, it appears extremely difficult to develop a strategy that is better at maximizing correct predictions than that provided by the seedings.

Given the finding that alternative strategies fail to improve upon the diagnosticity of seeding, and the fact that all correct selections have equal utility in terms of winning the pool, individuals should always predict the lower seeded teams to win according to a normative expected-utility model of choice under uncertainty (Einhorn, 1986; von Neumann & Morgenstern, 1947). A key question then is whether individuals largely rely on the normative

strategy of using the seedings of the teams in making their predictions, or whether they use other strategies instead. Assuming they do use other strategies, a second important question is whether individuals are able to successfully outperform a strategy relying only on the team seedings.

ESPN “Tournament Challenge” data

Figure 2 presents the first-round predictions of more than 3 million entries in the “Tournament Challenge” competition hosted by ESPN.com, a sports news website, for the years 2004 and 2005 (ESPN, 2005). The entry receiving the most points from correct predictions at the end of the tournament received a new car, and the runner-up a new television. As can be seen, the number of predicted upsets (i.e., the higher seed winning) in the first round climbs gradually to match the actual probability of such upsets.

Furthermore, it appears that individuals would have been better served, on average, by using a more conservative strategy. The average performance of participants in the Tournament Challenge in 2004 was 75.2% correct (2005: 72.9%). A strategy of always picking the lower seed would have resulted in 87.5% correct (2005: 75%). Thus, individuals on average predicted too many upsets and would have done better by always picking the lower seeds. A second question is whether individuals were better than chance at selecting these upsets. In other words, given the number of upsets picked, what score could be expected by chance alone? This expected score can be calculated by multiplying the percentage of predicted lower seed wins by the percentage of actual lower seed wins, and the percentage of predicted higher seed wins by the percentage of actual higher seed wins, and summing these two values. Individuals in the 2004 Tournament Challenge should have received 75.9% (2005: 71.6%) correct by chance alone. Thus, on average, performance was about what would be expected by chance, and was certainly worse than would have been attained by merely picking the lower seeds.

What might explain the poor performance of participants? One could argue that they were simply uninformed about college basketball, or were biased by their liking of the teams. Alternatively, participants could be using what they believe to be a rational strategy, basing their predictions on their own inaccurate perceptions of the strength of the teams. Finally, participants might simply be unable to predict upsets at a rate better than chance due to the fact that the teams have already been seeded by the experts organizing the tournament. Team statistics (e.g., win-loss record) are likely to be only weakly correlated with upsets, as the team seedings already control for this information. In other words, although the outcome of a basketball game is largely determined by the skill and preparation of the players, the seeding process in this tournament has “partialled-out” these non-random effects. As a result, the occurrence of an upset is due to factors that are extremely difficult to predict, and in this sense, random.

Assuming this is true, it raises the question of why participants would choose to predict so many upsets. One might consider the possibility that participants felt that they had no real chance of winning the grand prizes (literally less than one in a million), and therefore were either not motivated to be correct or believed they had to pick upsets in order to increase their chances. Indeed, in such a large competition, one may need near perfection to achieve the highest score, and so upset predictions have a higher utility (see also Kaplan & Garstka, 2001; Metrick, 1996). Individuals may also make predictions based upon team preferences, either because they are not motivated to be correct and hope their favorite teams will win, or because perceptions of team strength are biased. Although these are valid possibilities, we believe these predictions reflect a more fundamental pattern of “probability matching.”

Probability matching

Psychologists have long been interested in how individuals predict dichotomous outcomes under uncertainty, notably in the context of studies of probability learning (e.g., Grant, Hake,

& Hornscth, 1951; Humphreys, 1939; Lee, 1971). A typical paradigm involves two events, A and B, one of which randomly occurs on each trial of the experiment with a predetermined probability. Participants are asked to predict which outcome will obtain on each individual trial. From an expected utility viewpoint, as long as one of the events is more probable (i.e., both events do not occur equally often), one should always predict the more likely event (Grant et al., 1951; Lee, 1971; von Neumann & Morgenstern, 1947). However, in these studies, the probability of responses came to equal, or match, the probability of the event. Assuming that the outcomes are truly random or that one has no other diagnostic information, this behavior necessarily reduces one's chances of being correct (Einhorn, 1986). Consider the following example: Event A randomly occurs on 25% of all trials, and event B occurs on the other 75% of all trials. The participant is to predict whether A or B will occur on four separate trials. By always predicting event B, the participant can expect to be correct 75% of the time. Imagine, however, that the participant probability matches by predicting A on the first trial and B on the next three. There is only a 25% chance of being correct on this first trial, and a 75% of being correct on each of the other trials. Thus, the overall probability is $(1/4 \times .25) + (3/4 \times .75) = .625$.

A number of explanations have been offered for probability matching behavior, including participant's lack of understanding or failure to apply the notion of independence of events (Gal & Baron, 1996), boredom (Keren & Wagenaar, 1985), or increased utility of correctly predicting rare events (Tversky & Edwards, 1966). One particularly intriguing explanation for probability matching is that individuals do not understand or believe that the series of events is truly random and so they take a problem-solving approach to the task, rather than a statistical one (Braveman & Fischer, 1968; Goodnow, 1955; Morse & Runquist, 1960; Nies, 1962; Schul, Mayo, Burnstein, & Yahalon, 2007). That is, individuals attempt to develop a coherent causal explanation for the events in order to eliminate all error in prediction, rather

than trying to maximize their performance using a single diagnostic cue (Einhorn, 1986; Schul et al., 2007). Indeed, participants in these experiments often report hypotheses concerning patterns in the events or demonstrate sensitivity to recent patterns (e.g., Jarvik, 1951; Morse & Runquist, 1960). Instructions (Goodnow, 1955) or task features (Morse & Runquist, 1960; Peterson & Ulehla, 1965) that encourage a problem-solving mindset have been found to increase probability matching. For example, predicting an obviously random outcome, such as a roll of a die, reduces probability matching relative to a potentially pre-arranged outcome, such as the sequence of a stack of cards (Peterson & Ulehla, 1965). Because human behavior is seen as non-random, Schul et al. (2007) suggested that perceivers would also be more likely to take a problem-solving orientation when predicting events caused by human agency than when predicting events caused by a non-human agency. They asked participants to predict the color of a token placed in a matchbox with a color label. In two-thirds of the trials, the color on the box matched the color of the token inside. Individuals indeed demonstrated more probability matching when they thought a competitor had placed the tokens inside the box than when the tokens were supposedly randomly placed in the boxes.

Based on this past work, there is strong reason to believe that individuals demonstrate probability matching when predicting the outcomes of the NCAA basketball tournament because they believe the game outcomes are non-random. Individuals likely believe that the games are determined by the skill of the players, coaching decisions, training, and so on. As a result, they may come to think that they can correctly predict the outcome of all the games. The present studies were designed to examine this explanation for the prediction of upsets.

In addition, we sought to examine the consequences of probability matching for performance in tournament pools. Because the tournament teams have already been seeded by experts based on many of the same characteristics (e.g., win-loss record, strength of

opponents, etc.) individuals are likely to use, it should be extremely difficult to perform better than chance when picking upsets. Therefore, as with previous research using random events, participants are likely to suffer poorer performance on average as a result of probability matching rather than always predicting the lower seed to win.

Study 1 – Expert predictions

We conducted an initial study in which we examined how knowledgeable, motivated fans would make these predictions in a small pool of participants (thus increasing the chances of winning), and what the basis of these predictions might be. We examined whether upset predictions are a result of individuals liking the higher seeded team more or believing that the higher seeded team is actually better. Although we expected that these variables would influence predictions, we expected that a number of upsets would be picked despite the higher seed being seen as weaker and not being as well liked as the lower seed. Such predictions likely result from a belief that an occasional upset will occur and that overall score can be improved by selecting a few such upsets. We assume individuals consider past frequencies when making these upset predictions, leading to a probability matching phenomenon.

Method

Participants

The study took place at a large Midwestern university known for a high level of interest in college basketball. Men's basketball is by far the most popular varsity sport at the university. Participants were nineteen (six female, thirteen male) knowledgeable fans of NCAA men's basketball, ranging in age from 18 to 62. They were recruited through flyer advertisements for a study on how people make predictions for basketball games. The flyers were placed in and around the university campus the week before the tournament began. Participants indicated they were fans of college basketball and were able to correctly answer several relevant trivia questions (e.g., "Who is the coach of the Kentucky Wildcats?"). Thus, participants were

considered experts in the sense that they had a high level of interest in college basketball and demonstrated adequate knowledge on the trivia test. Participants were paid \$5 as compensation. In addition, the participant with the most number of accurate predictions received a \$20 prize.

Materials and procedure

Individuals interested in participating in the study completed the trivia questions and indicated whether they were fans of college basketball. From this initial pool, nineteen participants were identified. They were told that they would be asked to predict the outcome of all the games in the upcoming NCAA men's basketball tournament, and that the participant with the most number correct would receive a \$20 prize, in addition to the normal \$5 compensation. The study took place on Monday, March 13, 2000, the day after the schedule of games was announced and three days before the beginning of the tournament. Because of the small time frame in which to collect these data, the study was conducted on the internet using a website interface. Participants first completed a full tournament bracket, predicting the winner of all 63 games. They then rated to what extent they liked each team, using a scale ranging from 1 (not at all) to 7 (very much) and rated the strength of each team, using a scale ranging from 1 (very weak) to 7 (very strong). Following the completion of the tournament, participants received a debriefing as well as their compensation.

Results and Discussion

Overall frequency data

We focused our analyses on the 32 first round predictions. Figure 3 shows the overall frequency of upset predictions by team seeding. As can be seen, individuals demonstrated a probability matching pattern. The linear trend for the prediction data was significant, $F(1, 18) = 73.42, p < .001, \eta^2 = .80$. Thus, we observed the same basic probability matching pattern of tournament predictions among knowledgeable fans, each with a reasonable chance of winning

a prize, as was found in the ESPN Tournament Challenge data. Thus, probability matching in upset predictions does not seem to be a consequence of unmotivated or ill-informed decision-makers. One slight difference was that participants largely predicted wins by the lower seeds for the 1-4 seeded teams, where upsets are highly infrequent, and largely demonstrated probability matching for the remaining games, where upsets are more common. The smaller number of upset picks among the 1-4 seeded teams makes sense given that the frequency of such upsets is below 25%. Picking 1 upset out of 4 games thus exceeds the base-rate in these cases.

Liking and strength ratings

Five participants failed to complete all of the liking and strength measures, and so they were not included in analyses of these ratings. For each team, average ratings of liking and strength were calculated across participants. Liking ratings were generally near the mid-point of the scale and did not vary greatly across the 64 teams (overall $M = 4.00$, $SD = 0.41$). Strength ratings were also near the mid-point on average, but more variable across the 64 teams (overall $M = 3.90$, $SD = 1.33$).¹

Predicting choices

For each participant, we first determined the relative team liking and strength judgments for each selection by subtracting ratings of the higher seed from the ratings of the lower seed. We then created four categories of selections: cases in which the lower seeded team was rated equal to or higher than the higher seeded team in both liking and strength, in liking but not strength, in strength but not liking, or in neither. We then calculated the average frequency and proportion of selections (out of 32) of the lower seeded teams (no upset) and higher seeded team (upset) in each of these four categories. The results of this classification can be seen in Table 1. Not surprisingly, participants tended to rate the lower seeded teams as both stronger and more likeable, and were more likely to predict the lower seed would win.

However, it is also clear that selections were not based entirely on judgments of strength or greater liking of teams. In cases in which participants both liked the lower seed and agreed that this team was strong, the highest number of upset predictions was made. The proportion of upsets in this category significantly exceeded 0%, $t(13) = 5.98, p < .001, \eta^2 = .73$.

Conversely, even when they liked the lower seeded team less and believed this team was not as strong, participants predicted the lower seed to win at a rate significantly greater than 0%, $t(13) = 5.79, p < .001, \eta^2 = .72$, and twice as often as upsets. Thus, liking and strength judgments left a large number of picks unexplained.

To assess the relative contribution of liking and strength judgments in upset predictions, we calculated Goodman and Kruskal's (1954) λ statistic of predictive association. The λ statistic indicates the proportion of errors in explaining upset picks that can be eliminated by knowing the relative rating of the teams, ranging from 0 (no reduction in errors) to 1 (complete reduction in errors). Thus, we determined to what extent knowing relative liking and strength explained upset picks for each participant. In these analyses, knowing liking alone resulted in an average predictive association $\lambda = .09$, which did significantly differ from 0, $t(13) = 2.42, p < .05, \eta^2 = .31$. Thus, liking did contribute to the prediction of upsets, but only accounted for a small amount of variance in these selections. Knowing strength alone ($\lambda = .01, t(13) = 1.00, p = .34, \eta^2 = .07$) or both liking and strength ($\lambda = .12, t(13) = 1.95, p = .07, \eta^2 = .23$) did not significantly add to the prediction of upsets.² These findings support our view that participants largely pick upsets to increase their perceived chances of being correct rather than because they like the higher seeded team or believe it is actually a better team.

Assessments of accuracy

There were only three actual upsets in the first round of the 2000 tournament. Thus, prediction performance was relatively high. Participants actually scored an average of 27.8 (86.8%) correct, with scores ranging from 25 (78.1%) to 31 (96.9%). In comparison, the

strategy of taking the better seed yielded a score of 29 (90.6%). Participants thus did worse on average than simply taking the lower seed, $t(18) = 2.77, p < .05, \eta^2 = .30$.

A second question is whether individuals predicted the *correct* upsets. Therefore, we calculated expected score using the frequencies of upsets predicted by participants, multiplying the number of lower seed wins predicted by the number of actual lower seed wins and multiplying the number of higher seed wins predicted by the number of actual higher seed wins and adding these two values, resulting in an expected score of 27.1 (84.7%). Participants' performance (27.8, 86.8%) was not significantly higher than this score, $t(18) = 1.57, ns, \eta^2 = .12$. Thus, participants seem to be no better than the tournament organizers at predicting the first round outcomes of the tournament.

Study 2 – Event characteristics

It appears from the results of Study 1 that, despite their general lack of ability to do better than chance in predicting upsets, individuals nonetheless attempt to correctly predict them, resulting in probability matching. Furthermore, we found that ratings of liking and perceived team strength contributed little to the explanation of these upset predictions. One potential criticism of Study 1 is that the sample size was somewhat small, and thus the unique characteristics of the sample may have contributed to our findings. Specifically, one might argue that individuals were at a minimum picking specific upsets on the basis of liking of the teams, and perhaps our measures failed to fully capture this phenomenon. We therefore conducted a second study with novel teams and a larger sample to address these concerns. Because we believe these effects are due to the perception of basketball as a skilled event, similar effects should be found even among those unfamiliar with the tournament and the teams playing in it.

Study 2 also was designed to directly test the hypothesis that probability matching occurs because people believe basketball events are non-random and therefore predictable. We

compared predictions for a basketball tournament and a random computer simulation event, predicting that probability matching would be reduced when individuals perceive the events to be randomly determined (see also Schul et al., 2007). The structure of the tasks was otherwise identical. Furthermore, to rule out the possibility that the mere presentation of team information in the media is a strong cue that this information should be used (Grice, 1975), and to keep the two conditions as comparable as possible, participants were not given any additional information about the quality of the basketball teams. Because we propose that probability matching reflects a more fundamental belief that a basketball game is more predictable, probability matching should occur even when individuals do not have access to detailed information about the teams.

We assumed participants in the ESPN Tournament Challenge and the expert fans participating in Study 1 would be familiar with the frequency of upsets in past tournaments. To ensure participants in Study 2 were aware of these frequencies, they were provided with this information. In addition, recent research (e.g., Gigerenzer, 1994; Zhu & Gigerenzer, 2006) has shown that many apparent judgment biases are reduced when individuals are given more easily conceptualized frequency information. Given that we believe participants are using frequency information in an attempt to be 100% accurate, we believe the form of stimuli presentation will make little difference in probability matching effects. To provide additional evidence for the increased use of frequency rather than seeding information in the basketball condition, we also asked participants to indicate which strategy they were using to make their selections.

Finally, one could argue that individuals demonstrate probability matching in their upset predictions, not because they believe they can predict the games at a rate better than chance, but because they realize they need to predict more outcomes correctly than the other

participants (see also Metrick, 1996). If this were true, we would expect similar responses for predicting random events as for predicting basketball games.

Method

Participants

Participants were 40 students at a German university, randomly assigned to one of two task conditions. Participants completed the experiment along with several unrelated tasks, and were paid 5€ or received participation credit towards a course requirement as compensation. In addition, participants were told that the person with the most accurate predictions would receive an additional 15€ prize. This prize was actually awarded by a random lottery.

Task description manipulation

In the random condition, participants were told that the researchers were interested in how people predict the outcome of events. They were told that a computer had been programmed to make random selections from among pairs of unique events. There were supposedly 64 events, organized into four groups of eight pairs. The events varied in how likely they were to be selected by the computer, and that the options were ranked according to this probability, with 1 being the most likely and 16 the least likely. In addition, they were told that, although the exact probabilities could not be given to them, they would be told how often in 84 past simulations options with the same rank had been selected. An example was then given (see Appendix).

In the basketball condition, participants were told that the experimenters were interested in how people predict the outcome of sporting events. They were told that a tournament between 64 amateur teams would be taking place later in the summer in the United States, and that these teams varied in how good they were considered by experts. They were told that the teams were ranked according to their quality, with 1 being the best and 16 the worst. In addition, they were told that, while the exact probabilities of a team winning could not be

given to them, they would be told how often in 84 past games teams with the same rank had won. An example was then given (see Appendix).

Measures and procedure

Participants were randomly assigned to receive either the random simulation description or the basketball tournament description. They were given a bracket presenting all 32 games (or simulations), with a randomly generated three-letter moniker for each team (option) placed alongside the seeding and followed by the past frequency that similarly ranked teams (options) won out of 84 games (were selected by the computer out of 84 simulations). These values were taken from the actual frequency of upsets in the NCAA tournament for the years 1985-2005.

The participants then indicated on a separate sheet whether they had used a strategy of always picking the lower seed or always picking the seed more frequently winning (or selected). Participants then rated to what extent such events depend on chance or on the characteristics of the option/team using a 1 (completely depends on chance) to 7 (completely depends on characteristics) scale; to what extent they had relied on their knowledge or understanding of probability to make their predictions, using a 1 (more on my knowledge) to 7 (more on my understanding of probability) scale; and how motivated they were to win the 15€ prize, using a 1 (not at all) to 7 (very much) scale. At the completion of the study, participants received an email debriefing and the prize was awarded to a randomly selected participant.

Results and Discussion

Predictions

Figure 4 presents the selections for participants in the random simulation and basketball conditions by game seeding. A 2 (condition: basketball or random simulation) x 8 (seed: 1 vs. 16, 2 vs. 15, 3 vs. 14, 4 vs. 13, 5 vs. 12, 6 vs. 11, 7 vs. 10, 8 vs. 9) MANOVA using Wilks

criterion (see Maxwell & Delaney, 1990; Tabachnik & Fidell, 2001) was conducted on the number of upsets predicted, collapsed across the four groups. A planned contrast revealed that the Seed x Condition linear trend was significant, $F(1, 38) = 10.41, p < .01, \eta^2 = .22$. The linear trend was significant in the basketball condition, $F(1, 19) = 211.71, p < .001, \eta^2 = .92$, but not the random simulation condition, $F(1, 19) = 3.79, p < .07, \eta^2 = .17$. Participants demonstrated greater probability matching in the basketball condition than in the random simulation condition, consistent with our predictions. We thus found probability matching among individuals unfamiliar with the tournament using novel teams, effectively ruling out a liking-based explanation. Similarly, that the current findings replicated those of Study 1 (as well as the Tournament Challenge data) makes it extremely unlikely that the probability matching effects in Study 1 were due to unique sample characteristics.

Strategy

Self-reports of strategy use are presented in Table 2. Consistent with our predictions, participants in the basketball condition were more likely than those in the random condition to indicate that they picked the more frequent winners. Those in the random condition tended rather to use a strategy of selecting the lower seed, although this comparison was not significant. A subsequent analysis comparing the number of participants reporting using only the seedings (random condition $f = 5$, basketball $f = 1$) and those reporting using only the frequencies (random condition $f = 1$, basketball $f = 11$) also revealed a significant effect of condition (Fisher's Exact Test $p < .001, \phi = .75$). The results of the strategy measure thus support the conclusion that individuals predicting the basketball games were attempting to develop a strategy that would be 100% correct, relying on the frequency information more (and the seedings relatively less) than did those in the random simulation condition.

Ratings

Independent group t-tests were conducted on each of the ratings. Participants reported that the outcome of a basketball tournament game ($M = 5.10$) depends more on the characteristics of the team/option (and less on chance) than the outcome of a random simulation ($M = 2.70$), $t(38) = 5.30, p < .001, \eta^2 = .42$. There were no differences on the motivation measure or the item assessing whether individuals used their understanding of chance or general knowledge about the type of event in making their predictions ($ts < 1.38, ns, \eta^2 < .05$).

Mediation

We next tested for mediation of the moderating role of condition in the frequency of upset picks, a case of “mediated moderation” using the terminology of Muller, Judd, and Yzerbyt (2005). According to Muller et al. (2005), demonstrating this type of mediation requires first finding a treatment by moderator interaction on the dependent variable. As discussed, the Seed x Condition linear trend was indeed significant. Next, an effect of the moderator should be observed on the proposed mediator. This effect was observed for the extent to which the event depends on chance or the characteristics of the team or option (which we will refer to here as “Option characteristics”). Third, a significant treatment by mediator interaction must be observed on the dependent variable, accompanied by the reduction of the treatment by moderator interaction to non-significance (Muller et al., 2005). That is, the Seed x Option characteristics linear trend must be significant, accompanied by the reduction of the Seed x Condition linear trend to non-significance.

The Option characteristics rating was therefore included as a covariate in a separate MANCOVA model, with seed and condition as independent variables. In this analysis, the Seed x Option characteristics linear trend was significant, $F(1, 37) = 4.19, p < .05, \eta^2 = .10$, and inclusion of this rating reduced the Seed x Condition linear effect to nonsignificance, $F(1,37) = 1.47, ns, \eta^2 = .04$. Thus, there was evidence for mediation of the moderating role of

condition by ratings of the extent to which the event depends on chance or the characteristics of the team or option.

To explore this finding, three groups were created from the Option characteristics rating, with one group believing the event depended more on chance (i.e., scores ranging from 1 to 3), another believing the event depended more on the characteristics of the team/option (i.e., scores of 6 or 7), and a third group in between (i.e., scores of 4 or 5). Separate MANOVAs on the number of upset predictions with seeding as the independent variable were conducted for each group. Among those believing most strongly that the event depended on the characteristics of the team/option, a strong linear trend was observed, $F(1, 10) = 50.38, p < .001, \eta^2 = .83$, see Figure 5. The linear trend was also significant for those moderately believing the event depended on the characteristics of the team/option, $F(1, 13) = 79.82, p < .001, \eta^2 = .86$. However, among those believing the outcome depended on chance, no linear trend was found, $F(1, 14) < 1, p > .42, \eta^2 = .05$. Thus, the present study demonstrates that probability matching occurs in predicting the outcomes of this type of basketball tournament because individuals believe the game outcomes are non-random and therefore predictable. This finding supports past research showing that events that are due to human factors are perceived as more predictable, and therefore individuals attempt to achieve 100% correctness by predicting the less likely outcome at the rate of past frequency of occurrence (Goodnow, 1955; Grant et al., 1951; Schul et al., 2007).

Study 3 – Does more information matter?

One unexpected aspect of Study 2 was the relatively high number of upset predictions in the random condition. Although participants in this condition did not demonstrate probability matching, it was also the case that responses in this condition were not completely normative. It is possible that participants did not understand the computer simulation instructions or took “random” to mean the probabilities of an upset were highly variable, rather than that the

outcomes were randomly determined with a given probability. In this case, participants would be likely to respond haphazardly rather than making a prediction based on a consideration of the likelihood of a given outcome. More closely examining the individual responses, four participants in this condition did rely entirely on the seedings, whereas none in the basketball tournament condition did so. On the other hand, there was also evidence that a number of individuals may have misunderstood the random simulation instructions to mean that the probabilities were highly variable. For example, nine participants in the random condition predicted at least one upset in the 1-16 selection, in which no upset had ever occurred. Only one participant in the basketball condition did so. Thus, it appears that the random simulation instructions led some participants to respond normatively and others to respond haphazardly. Study 3 was designed eliminate this potential problem.

A second possible alternative explanation for the results of the previous studies involves the fact that participants in the basketball condition were not given more detailed information about the basketball teams. This was intentionally done to ensure that there was no experimental demand for participants to use this information and pick upsets, and to keep the random and basketball conditions in Study 2 as comparable as possible. However, one could argue that probability matching occurred in these studies because participants did not have enough information about the quality of the teams and that they would pick the lower seeds when provided with this information. Although we recruited knowledgeable participants in Study 1, we cannot be certain they possessed information about the team's record, shooting percentage, etc., when making their predictions. Therefore, we included a condition in which information about the team's performance during the season was also provided.

Method

Participants

Thirty-four German university students participated in return for payment of 2.50€ or a half-hour of research participation credit. They were told that the participant with the most number of correct predictions at the conclusion of the study would receive an additional 15€ prize. This prize was actually determined randomly. Participants were randomly assigned to one of the three experimental conditions.

Materials and procedure

As in Study 2, participants were told they would be involved in a study on making predictions and received one of the three task descriptions.

Random drawing condition

In the random drawing condition, participants were told that the experimenters had designed a lottery in which three-letter “options” were written on slips of paper and placed in a box. There were always two different options in the box, and they were asked to predict which option would be drawn from the box. They were also told that some options had more slips of paper in the box than others; specifically, those options with a rank of 16 had relatively the fewest slips of paper and those with rank 1 had relatively the most slips of paper in the box. Furthermore, the lottery was such that rank 1 options were placed in the same box with rank 16 options, rank 2 options with rank 15 options, and so on. They were told the experimenters could not tell them the exact number of slips of paper for each of the two options that were placed in a box, but that they would be told the number of times each option had been selected in 84 previous drawings. They were provided with an example (see Appendix).

Basketball condition

The basketball condition was identical to that described in Study 2.

Basketball plus information condition

The basketball plus information condition was identical to the basketball condition, but participants were also provided with information about how many times the team had won and lost in that season, what percentage of their shots at the basket were successful (i.e., offense), what percentage of their opponents' shots at the basket were successful (i.e., defense), and how difficult their schedule was compared to teams in the rest of the country. This information was taken from the actual teams in the 2005 NCAA tournament. They were provided with an example (see Appendix).

Dependent measures

The brackets were identical across conditions, with the exception that in the basketball plus information condition the team statistics were provided along with the frequency information. After completing their bracket, participants indicated which strategy (seeding or frequency) they had used in completing the task, and completed the same ratings as in Study 2.

Results and Discussion

Predictions

Figure 6 presents the selections for participants in the random drawing, basketball, and basketball plus information conditions by game seeding. A 3 (condition: basketball, basketball plus information, or random) x 8 (seed: 1 vs. 16, 2 vs. 15, 3 vs. 14, 4 vs. 13, 5 vs. 12, 6 vs. 11, 7 vs. 10, 8 vs. 9) MANOVA was conducted on the number of upsets predicted, collapsed across the four groups. The predicted Seed x Condition linear trend emerged, $F(2, 31) = 4.73, p < .05, \eta^2 = .23$. The linear trend was significant in the basketball condition, $F(1, 11) = 54.21, p < .001, \eta^2 = .83$, and the basketball plus information condition, $F(1, 10) = 32.38, p < .001, \eta^2 = .76$, but not the random drawing condition, $F(1, 10) = 2.85, p = .12, \eta^2 = .22$. Replicating Study 2, participants demonstrated greater probability matching in the two basketball conditions than in the random drawing condition. No differences emerged between

the two basketball conditions, all F s < 1 , ns, $\eta^2 < .03$. Participants in the random condition picked fewer upsets than in Study 2 (thus responding more normatively), as predicted. These results suggest that the lottery instructions were more easily understood than the computer simulation instructions. However, this procedural change did not increase the degree of probability matching in this condition, suggesting that reduced probability matching in the random condition in Study 2 cannot be attributed to a lack of understanding. We also replicated the effects in the basketball condition when providing individuals with more information about the teams. Therefore, it is also not the case that more knowledge about the quality of the teams led individuals to always select the better team.

Strategy

There were no differences between the two basketball conditions in self-reports of strategy (Fisher's Exact Test p s $> .24$, $\phi < .24$). These two conditions were then combined to test for the hypothesized differences with the random condition, see Table 3. Participants in the basketball conditions were less likely to report selecting the lower seed, relative to those in the control condition. As in Study 2, those in the basketball conditions tended to report using a strategy of picking the more frequent winners more than did those in the control condition, although this comparison did not reach significance. A subsequent analysis comparing the number of participants reporting using only the seedings (random condition $f = 5$, basketball and basketball plus information conditions $f = 1$) and those using only the frequencies (random condition $f = 1$, basketball and basketball plus information conditions $f = 6$) also revealed a significant effect of condition (Fisher's Exact Test $p = .03$, $\phi = .69$). Replicating the results of Study 2, those in the basketball conditions were more sensitive to frequency information.

Ratings

A planned-comparison indicated that participants used their understanding of probability more in the random condition ($M = 5.32$) than in the basketball condition ($M = 4.25$) and in the basketball plus information condition ($M = 4.00$), $t(31) = 2.46$, $p < .05$, $\eta^2 = .16$. There was no difference between the two basketball conditions, $t(31) < 1$, ns, $\eta^2 = .01$. There were no other effects on the other rating measures ($F_s < 1$, ns, $\eta^2 < .06$).

Mediation

The significant Seed x Condition linear trend satisfied the first step of the test of mediated moderation (Muller et al., 2005). The condition effect on the rating of the extent to which participants used their knowledge of the event or understanding of probability (which we will refer to here as “Knowledge”) satisfied the second step. Next, Knowledge was included as a covariate in a MANCOVA with seed and condition as independent variables. The Seed x Knowledge linear trend was significant, $F(1, 30) = 4.28$, $p < .05$, $\eta^2 = .12$, and inclusion of this variable reduced the Seed x Condition linear effect to nonsignificance, $F(2, 30) = 2.46$, $p = .10$, $\eta^2 = .14$, evidence for mediation of the moderating role of condition on upset predictions (Muller et al., 2005).

To explore this finding, three groups were created from the Knowledge rating, with one group indicating that they used their knowledge of the event when making their predictions (i.e., scores ranging from 1 to 3), another indicating they used their understanding of probability (i.e., scores ranging from 5 to 7) and a third group in between (i.e., scores of 4). Separate MANOVAs on the number of upset predictions with seeding as the independent variable were conducted for each group. Among the group indicating most strongly that they had used their knowledge of the event, a strong linear trend was observed, $F(1, 10) = 86.30$, $p < .001$, $\eta^2 = .90$, see Figure 6. The linear trend was also significant for those rating their use of knowledge about the team/option as only moderate, $F(1, 11) = 21.27$, $p < .001$, $\eta^2 = .66$. However, among those indicating that they had used their understanding of probability, no

effect was found, $F(1, 10) = 3.76, p > .08, \eta^2 = .27$. Thus, the results demonstrated that individuals predicting a basketball tournament are likely to rely on their knowledge of basketball in making predictions, rather than normative strategies of probability.³ As a result, they tend to consider and make predictions consistent with past frequencies rather than relying solely on the seedings.

General Discussion

We began by investigating how individuals perform in one of the most the popular forms of sports gambling in the U.S., NCAA basketball tournament pools. We were interested in how individuals make these predictions (in particular, why they pick upsets) and whether they are better than chance at predicting the outcomes of these games, questions that have received little or no scientific attention to this point.

Upset predictions

The present studies demonstrated that individuals are likely to make upset predictions in a bid to be 100% correct. Participants in the basketball conditions specifically indicated that they were using past frequency of upsets to guide them in their selections, resulting in probability matching behavior. That is, individuals picked upsets at a rate approaching historical frequency. Furthermore, individuals demonstrated this behavior because they believe that a basketball game is a skilled, non-random event and is therefore predictable. When they thought that they would be predicting a random drawing or computer simulation, probability matching was reduced and individuals indicated they were relying on the seeding in making their selections. Moreover, mediation analyses revealed that these differences were explained by the extent to which individuals believed the event was determined by properties of the teams (rather than chance) and therefore used their knowledge of the event in making their prediction (rather than their understanding of probability). These findings replicate past research showing that probability matching is stronger for events that are seen as non-random

or due to human agency (Braveman & Fischer, 1968; Goodnow, 1955; Grant et al., 1951; Morse & Runquist, 1960; Nies, 1962; Schul et al., 2007). Thus, our results indicate that upset selections in predicting the NCAA tournament reflect at least in part a classic probability learning phenomenon.

We also ruled out a number of alternative explanations for our effects. Although liking the higher seed seemed to be predictive of an increased rate of upset selections, it was not the case that most upsets involved picking a well-liked “underdog.” Indeed, the most frequent type of upset prediction concerned teams that were rated both weaker and less likable than the lower seed. We also observed the same basic pattern of upset predictions among experts and those unfamiliar with the tournament, and with individuals predicting the outcomes of an actual tournament or a hypothetical one with randomly generated team monikers. Thus, while individuals may justify certain upset predictions on the basis of liking the team, this process does not seem to explain the selection of upsets more generally. Furthermore, we observed robust probability matching for upset predictions across three different samples, and showed that similar effects are observed in the ESPN Tournament Challenge with well over a million entries per year. It is therefore highly unlikely that our effects were in some way due to the unique characteristics of our samples.

A probability matching effect occurred regardless of whether participants had additional information about the teams, nor was it the case that they were due to differences in motivation or because picking the less frequent event had a higher utility (Metrick, 1996; Tversky & Edwards, 1966). Motivation was equally high for participants in the random selection conditions of Studies 2 and 3, and participants in all studies had a relatively strong incentive to be accurate. Furthermore, if probability matching were only a matter of increased utility of certain predictions, the type of event (i.e., whether random selections or basketball games) should have no effect on the degree of probability matching demonstrated. Finally, it

is unlikely that participants demonstrated probability matching as a result of boredom. The task was relatively short compared to most probability learning experiments in which hundreds or even thousands of trials are presented, and the task was largely the same across conditions.

Accuracy

A further finding was that, on average, individuals predicted fewer games correctly as a result of probability matching behavior, relative to the normative strategy of always selecting the lower seed. Probability matching can only lead to better performance if one can exceed chance in predicting the less likely event. Our participants appeared unable to do so. Indeed, past research has indicated it is extremely difficult to surpass the seedings in maximizing the number of correct predictions (Kaplan & Garstka, 2001; Lane & Damiano, 1994). Similar findings are reported by Cantinotti, Ladouceur, and Jacques (2004), who found that sports gamblers may be able to correctly predict the outcomes of games, but that they are unable to outperform the odds set by bookkeepers. As discussed, in a competitive gambling pool, there is some utility to picking the occasional upset; otherwise, everyone would receive the same score (Kaplan & Garstka, 2001; Metrick, 1996). However, at least in the typical small office pool where the chances of winning are higher, participants would likely be better served predicting a limited number of upsets.

Although the current research did not directly compare the performance of experts and non-experts, the findings suggest that non-experts would likely perform as well as experts in these pools. There is at least anecdotal evidence that individuals with little basketball knowledge or even those using seemingly nonsensical strategies occasionally outperform those making predictions based upon team statistics (e.g., Clarke, 2005; "Q&A Jacob Dodson," 2007). If novices have access to the seeding information, they have direct access to the most useful strategy available, making it likely that they will perform quite well. In

contrast, experts may continue to rely on cues (e.g., team record, schedule strength, etc.) that do not improve prediction over and above the tournament seeding, and considering this information may lead them to stray from this optimal strategy. Research by Gigerenzer and colleagues demonstrates that simple heuristics can often outperform more complex processes requiring the use of extensive data (e.g., Gigerenzer & Goldstein, 1996; Goldstein & Gigerenzer, 1999). Indeed, these “less is more” effects show that the use of too many cues can undermine judgment quality. Similarly, judgment quality can be undermined by extensive reasoning and deliberation (Dijksterhuis, Bos, Nordgren, & van Baaren, 2006; Dijksterhuis & van Olden, 2006; Wilson et al., 1993; Wilson & Schooler, 1991).

Direct evidence for this perspective is provided by a pair of studies conducted by Halberstadt and Levine (1999), in which individuals predicted third round games in the NCAA men’s tournament. Half were asked to provide reasons for their selections and the other half were asked to use their “gut feeling.” Those induced to give reasons for their predictions were less accurate and deviated more from the predictions of expert odds makers, relative to those not asked to give their reasons. Thus, in the context of predicting the NCAA tournament, less reasoning involving additional information may often be better. The use of a simple heuristic such as tournament seeding is as good as or better than any other strategy. Future research should therefore directly examine the role of expertise in predicting NCAA tournament games, in particular whether experts are more likely than novices to over-rely on reasoning and team-statistics, resulting in “less is more” effects.

Implications for probability matching research

The present work focuses on a specific type of judgment, predicting NCAA tournament games, yet judgment behavior was explained by a basic cognitive phenomenon. We believe the current findings therefore have implications for a variety of other domains in which individuals must make repeated predictions under uncertainty. The present studies add to the

probability matching literature in two important ways. Along with the findings of Schul et al. (2007), these studies provide one of the first demonstrations that probability matching is stronger for predictions of social behavior than for predictions of random events. The present research also extends these findings by demonstrating that beliefs about the characteristics of the events and their predictability mediate this difference. The vast majority of probability matching studies have examined relatively contrived tasks such as predicting the appearance of a light or the roll of dice (cf., Goodnow, 1955; Grant et al., 1951; Lee, 1971). However, lay theories hold that the behavior of individuals is due to coherent personality structures which can be discovered by perceivers (see Asch, 1946; Sherman & Hamilton, 1996). Thus, an intriguing question for future research is whether probability matching would be found for behavioral predictions on the basis of trait or stereotype expectations and whether the strength of these expectations might moderate these effects.

References

- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 258-290.
- Braveman, N. S., & Fischer, G. J. (1968). Instructionally induced strategy and sequential information in probability learning. *Journal of Experimental Psychology, 4*, 674-676.
- Cantinotti, M., Ladouceur, R., & Jacques, C. (2004). Sports betting: Can gamblers beat randomness? *Psychology of Addictive Behaviors, 18*, 143-147.
- Clarke, L. (2005, March 19). March Madness turns gamblers into basket cases. *Washington Post*, p. A1.
- Dijksterhuis, A., Bos, M., Nordgren, L. F., & van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science, 311*, 1005-1007.
- Dijksterhuis, A., & van Olden, Z. (2006). On the benefits of thinking unconsciously: Unconscious thought can increase post-choice satisfaction. *Journal of Experimental Social Psychology, 42*, 627-631.
- Einhorn, H. J. (1986). Accepting error to make less error. *Journal of Personality Assessment, 50*, 387-395.
- ESPN. (2005). *Tournament Challenge national bracket*. Retrieved May 5, 2005, from <http://games.espn.go.com/tcm/html/nationalBracket>.
- Gal, I., & Baron, J. (1996). Understanding repeated simple choices. *Thinking and Reasoning, 2*, 81-98.
- Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 129-161). Oxford, England: John Wiley & Sons.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650-669.

- Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer, P. M. Todd & A. R. Group (Eds.), *Simple heuristics that make us smart* (pp. 37-58). New York: Oxford University.
- Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association, 49*, 732-764.
- Goodnow, J. J. (1955). Determinants of choice distribution in two-choice situations. *American Journal of Psychology, 68*, 106-116.
- Grant, D. A., Hake, H. W., & Hornseth, J. P. (1951). Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *Journal of Experimental Psychology, 42*, 1-5.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41-58). New York: Academic Press.
- Halberstadt, J. B., & Levine, G. M. (1999). Effects of reasons analysis on the accuracy of predicting basketball games. *Journal of Applied Social Psychology, 29*, 517-530.
- Humphreys, L. G. (1939). Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *Journal of Experimental Psychology, 25*, 294-301.
- Jarvik, M. E. (1951). Probability learning and a negative recency effect in serial anticipation of alternative symbols. *Journal of Experimental Psychology, 41*, 291-297.
- Kaplan, E. H., & Garstka, S. J. (2001). March madness and the office pool. *Management Science, 47*, 369-382.
- Keren, G., & Wagenaar, W. A. (1985). On the psychology of playing blackjack: Normative and descriptive considerations with implications for decision theory. *Journal of Experimental Psychology: General, 114*, 133-158.
- Lane, I. M., & Damiano, P. L. (1994). Determining decision-making effectiveness using NCAA basketball tournament results. *Journal of Sport Behavior, 17*, 79-86.

- Lee, W. (1971). *Decision theory and human behavior*. New York: John Wiley & Sons.
- Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analyzing data: A model comparison perspective*. Belmont, CA: Wadsworth/Thompson Learning.
- Metrick, A. (1996). March madness? Strategic behavior in NCAA basketball tournament betting pools. *Journal of Economic Behavior and Organisation*, 30, 159-172.
- Morse, E. B., & Runquist, W. N. (1960). Probability-matching with an unscheduled random sequence. *American Journal of Psychology*, 73, 603-607.
- Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology*, 89, 852-863.
- Nies, R. C. (1962). Effects of probable outcome information on two-choice learning. *Journal of Experimental Psychology*, 64, 430-433.
- Peterson, C. R., & Ulehla, Z. J. (1965). Sequential patterns and maximizing. *Journal of Experimental Psychology*, 69, 1-4.
- Q&A Jacob Dodson: Facebook's \$25K NCAA tourney pool winner opens up. (2007, April 10). *SIO on Campus*, 2007.
- Schul, Y., Mayo, R., Burnstein, E., & Yahalon, N. (2007). How people cope with uncertainty due to chance or deception. *Journal of Experimental Social Psychology*, 43, 91-103.
- Sherman, S. J., & Hamilton, D. L. (1996). Perceiving persons and groups. *Psychological Review*, 103, 336-355.
- Tabachnik, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). New York: HarperCollins.
- Tversky, A., & Edwards, W. (1966). Information versus reward in binary choices. *Journal of Experimental Psychology*, 71, 680-683.
- von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. New York: Princeton.

- Wilson, T. D., Lisle, D., Schooler, J. W., Hodges, S. D., Klaaren, K. J., & LaFleur, S. J. (1993). Introspecting about reasons can reduce post-choice satisfaction. *Personality and Social Psychology Bulletin*, *19*, 331-339.
- Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, *60*, 181-192.
- Zhu, L., & Gigerenzer, G. (2006). Children can solve Bayesian problems: The role of representation in mental computation. *Cognition*, *98*, 287-308.

Author Note

Sean M. McCrea, Department of Psychology, University of Konstanz.

Edward R. Hirt, Department of Psychological and Brain Science, Indiana University
Bloomington.

The authors wish to thank Aaron Coleman for designing the internet interface for Study 1, and Judith Stöckel who collected data for the second and third studies. Correspondence concerning this article should be addressed to Sean M. McCrea, Fachbereich Psychologie, Universität Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany. Electronic mail may be sent via the Internet to Sean.McCrea@uni-konstanz.de

Appendix

Study 2, Random condition

1 ANA (84/84) 16 APP (0/84)

In this example, you would be asked to choose between Option ANA, which has a rank of 1, and Option APP, which has a rank of 16. Options with a rank of 1 were selected in 84 out of 84 past simulations, whereas options with a rank of 16 were selected in 0 out of 84 past simulations. This does not mean that an option with a rank of 16 *will never* be selected, but that to this point, it never has happened. You should then select which option will be chosen by the computer this time.

Study 2, Basketball condition

1 ANA (84/84) 16 APP (0/84)

In this example, you would be asked to choose between team ANA, which has a rank of 1, and team APP, which has a rank of 16. Teams with a rank of 1 won in 84 out of 84 past games, whereas teams with a rank of 16 won in 0 out of 84 past games. This does not mean that a team with a rank of 16 *will never* win, but that to this point, it never has happened. You should then select which team will win this time.

Study 3, Random condition

1: ANA (84/84) 16: APP (0/84)

In this example, you would be asked to choose between Option ANA, which has a rank of 1, and Option APP, which has a rank of 16. Thus, there are many more ANA slips of paper than APP slips of paper. In addition, options with a rank of 1 have been selected in 84 out of 84 past lotteries, whereas options with a rank of 16 have been selected in 0 out of 84 past lotteries. This does not mean that an option with a rank of 16 *will never* be drawn in the

lottery, but that to this point, it never has happened. You would then select which option you believe will be drawn this time in the lottery.

Study 3, Basketball plus information condition

1: ANA (84/84)	16: APP (0/84)
Wins-Losses: 30-1	Wins-Losses: 25-5
Opponent ranking: 6	Opponent ranking: 116
Own: 51%	Own: 45%
Opponent: 35%	Opponent: 40%

In this example, you would be asked to choose between Team ANA, which has a rank of 1, and Team APP, which has a rank of 16. Thus, team ANA is considered much better than team APP. In addition, teams with a rank of 1 have won in 84 out of 84 past games, whereas teams with a rank of 16 have won in 0 out of 84 past games. This does not mean that a team with a rank of 16 *will never* win in the tournament, but that to this point, it never has happened.

Team ANA has won 30/31 of its games this season, played the 6th most difficult schedule in the country, scored on 51% of their shots at the basket and allowed their opponents to score on 35% of their shots. Team APP has won 25/30 of its games this season, played only the 116th most difficult schedule in the country, scored on 45% of their shots at the basket and allowed their opponents to score on 40% of their shots.

Footnotes

¹ Although rated most positive overall by participants on the liking measure ($M = 5.21$), the university's team was seeded sixth in the tournament. Thus, liking of this team was not a source of upset predictions, as it was already favored to win. The least liked team was one of the university's arch-rivals ($M = 3.14$), seeded fifth. Nonetheless, only one participant predicted this rival team would be upset, exactly the same number that predicted the university's team would be upset. Thus, it does not appear that the location of the study played a significant role in the results reported here, either through undue home team support or "Schadenfreude" for a hated rival. Furthermore, participants were unlikely to have strong allegiances to the vast majority of the other teams participating in the tournament, and the outcome of the other 31 games had no direct consequence for the chances their favorite team would win.

² We also calculated the relative proportion of predicted upsets (i.e., the number of predicted upsets out of the total number of selections in each category) in each of the four categories for each participant. Analyses on this measure also revealed significant effects of liking, but not knowledge.

³ The reader may ask why Option characteristics served as the mediator in Study 2 but Knowledge served as the mediator in Study 3. The most likely explanation involves the change in the description in the random condition. Selection in the random lottery (Study 3) was generally seen as more due to the characteristics of the event, and less to chance, than was selection in the computer simulation (Study 2). In contrast, ratings in the basketball conditions were largely consistent across the two studies.

Table 1.

Proportion of lower and higher seed selections by relative liking and strength ratings (Study

1)

	<u>Like lower seed</u>		<u>Like higher seed</u>	
	Lower seed stronger	Higher seed stronger	Lower seed stronger	Higher seed stronger
Picked lower seed				
Mean <i>f</i>	13.58	1.68	3.84	1.16
Mean proportion (out of 32 selections)	.58	.07	.16	.05
Picked higher seed				
Mean <i>f</i>	1.42	0.26	1.05	0.58
Mean proportion (out of 32 selections)	.06	.01	.04	.02
Mean relative proportion of upsets within category	.09	.09	.22	.30

Table 2.

Frequency of reported strategy use (Study 2)

Condition	Picked higher seed	Picked more frequent winner	n
Random	7	3	20
Basketball	3	14	20
Total	10	17	40
Fisher's Exact Test (p)	.14	<.001	
ϕ	.23	.56	

Table 3.

Frequency of reported strategy use (Study 3)

Condition	Picked higher seed	Picked more frequent winner	n
Random	6	2	11
Basketball / Basketball plus information	4	9	23
Total	10	11	34
Fisher's Exact Test (p)	.04	.21	
ϕ	.38	.21	

Figure Caption

Figure 1. Percentage of actual upsets in NCAA tournament by seeding for years 1985-2005

Figure 2. Percentage of predicted upsets by higher seed from ESPN 2004 and 2005

Tournament Challenge Contest

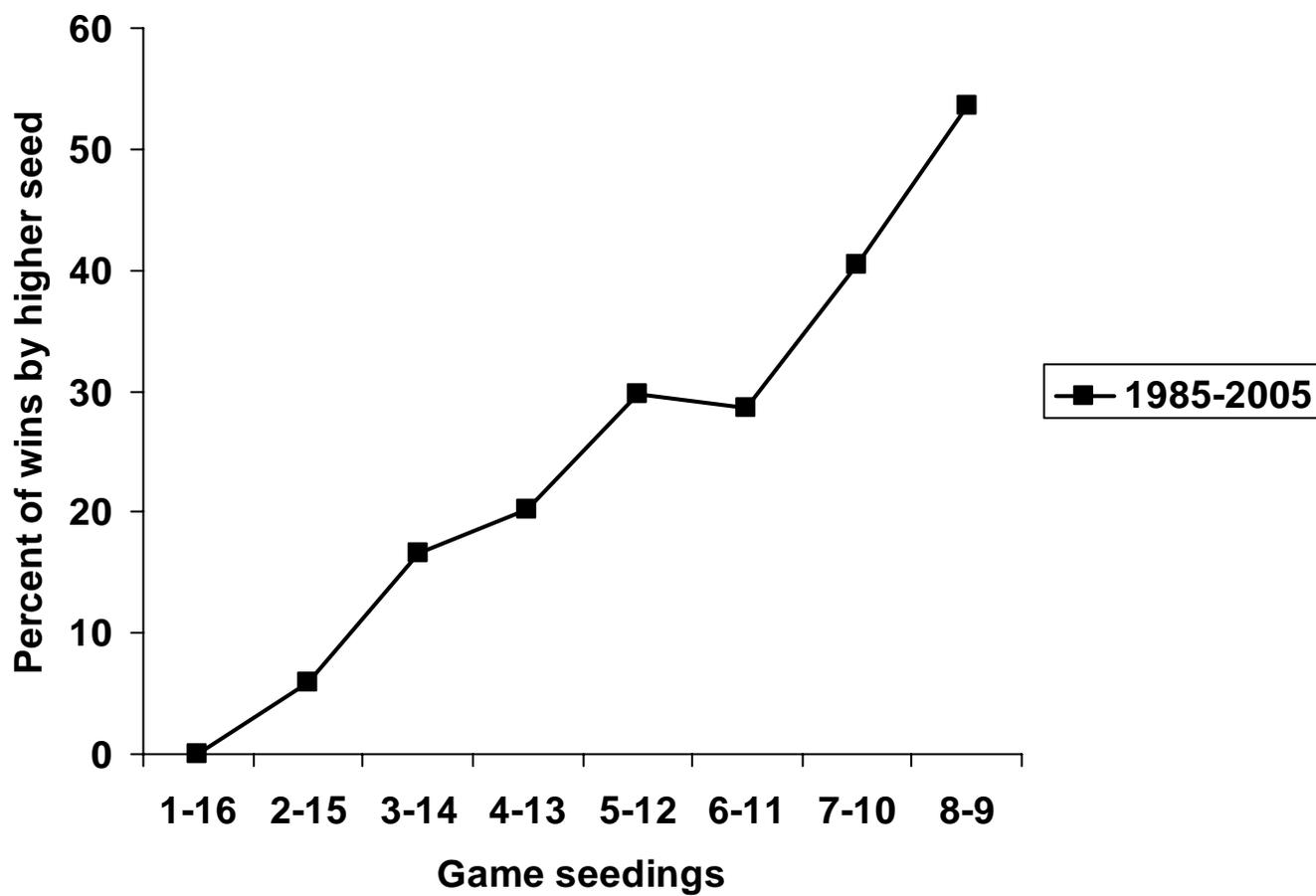
Figure 3. Percentage of predicted upsets by higher seed (Study 1)

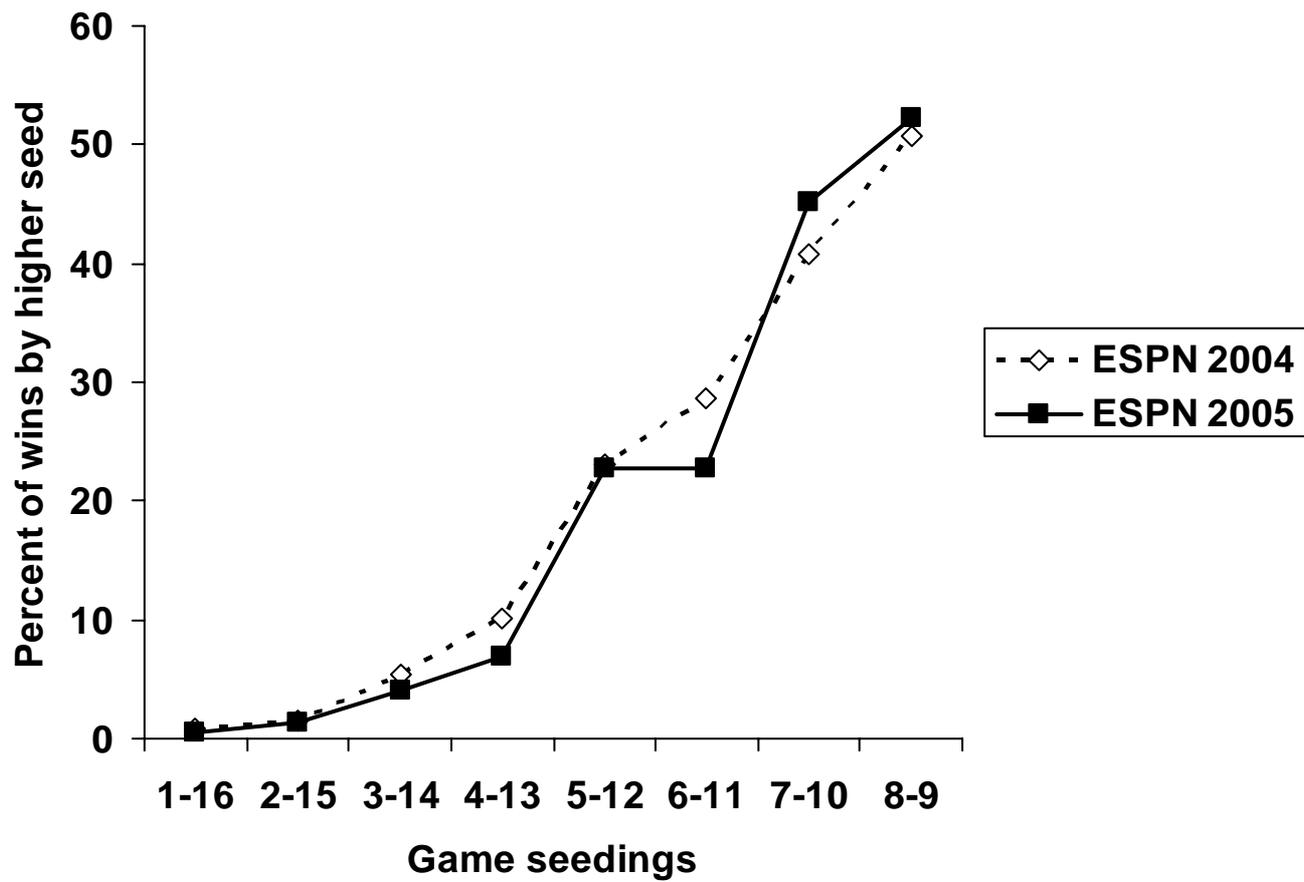
Figure 4. Percentage of predicted upsets by higher seed by condition (Study 2)

Figure 5. Percentage of predicted upsets by higher seed by option characteristics rating (Study 2)

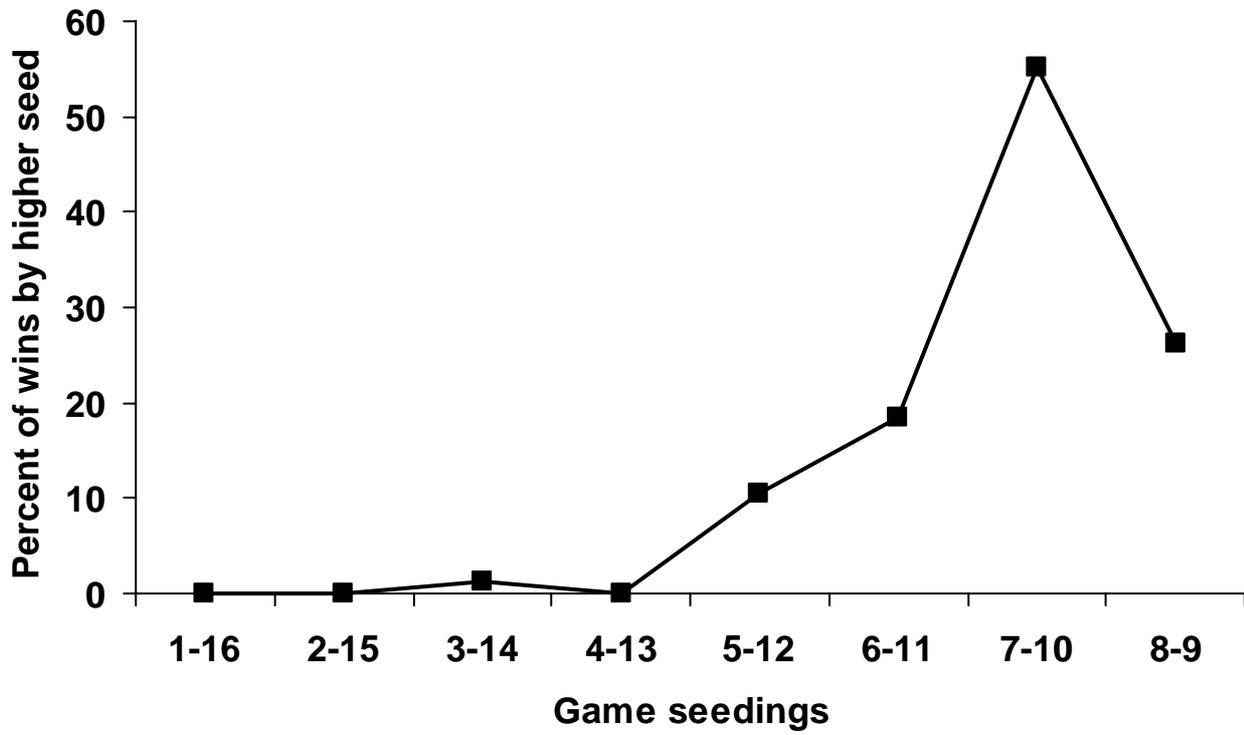
Figure 6. Percentage of predicted upsets by higher seed by condition (Study 3)

Figure 7. Percentage of predicted upsets by higher seed by knowledge rating (Study 3)

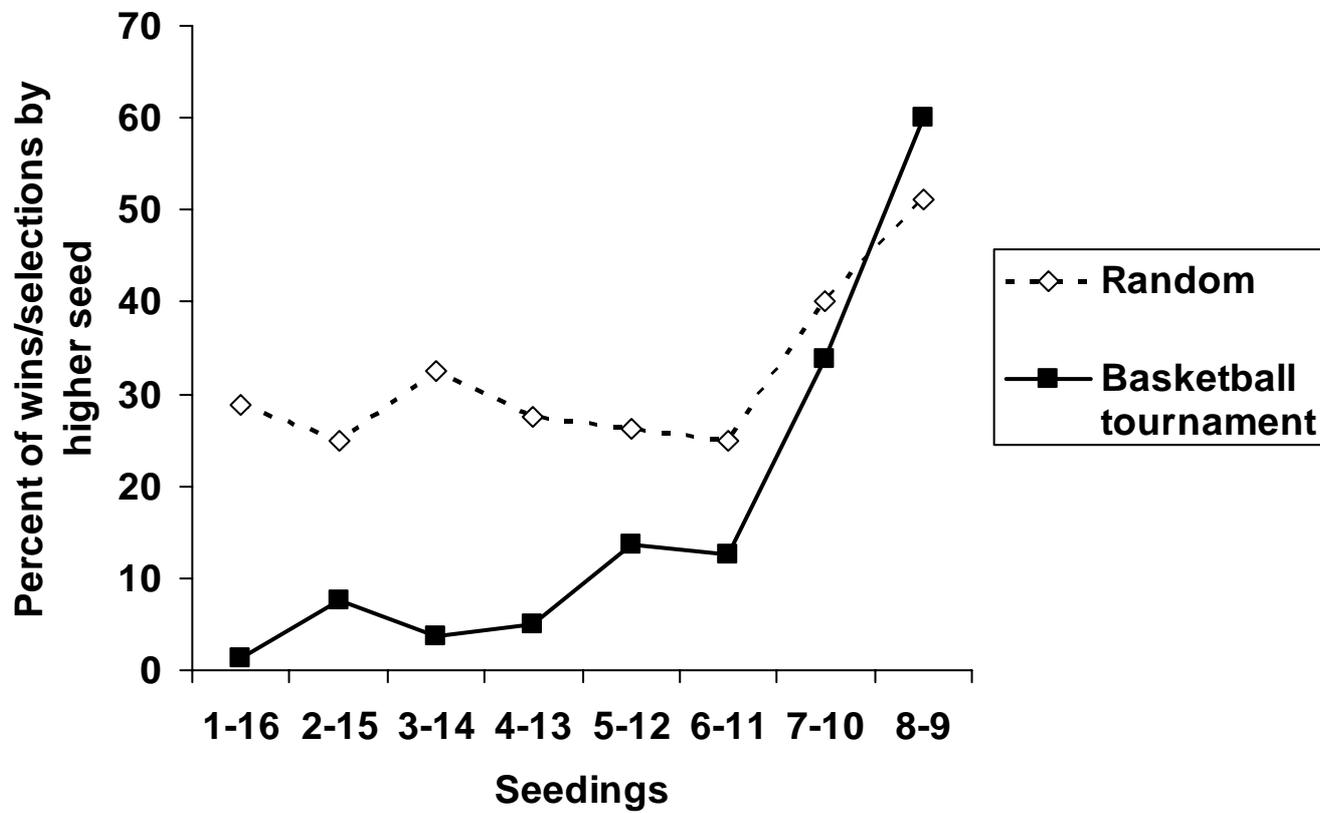




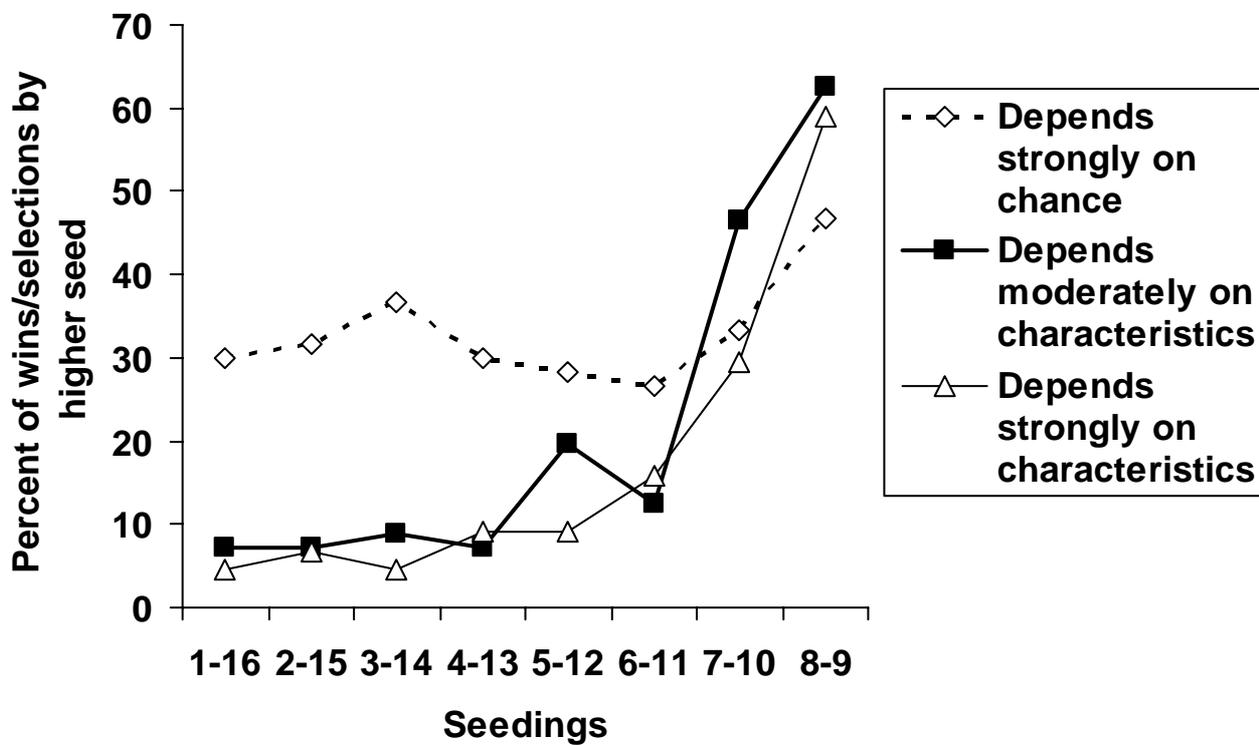
Study 1



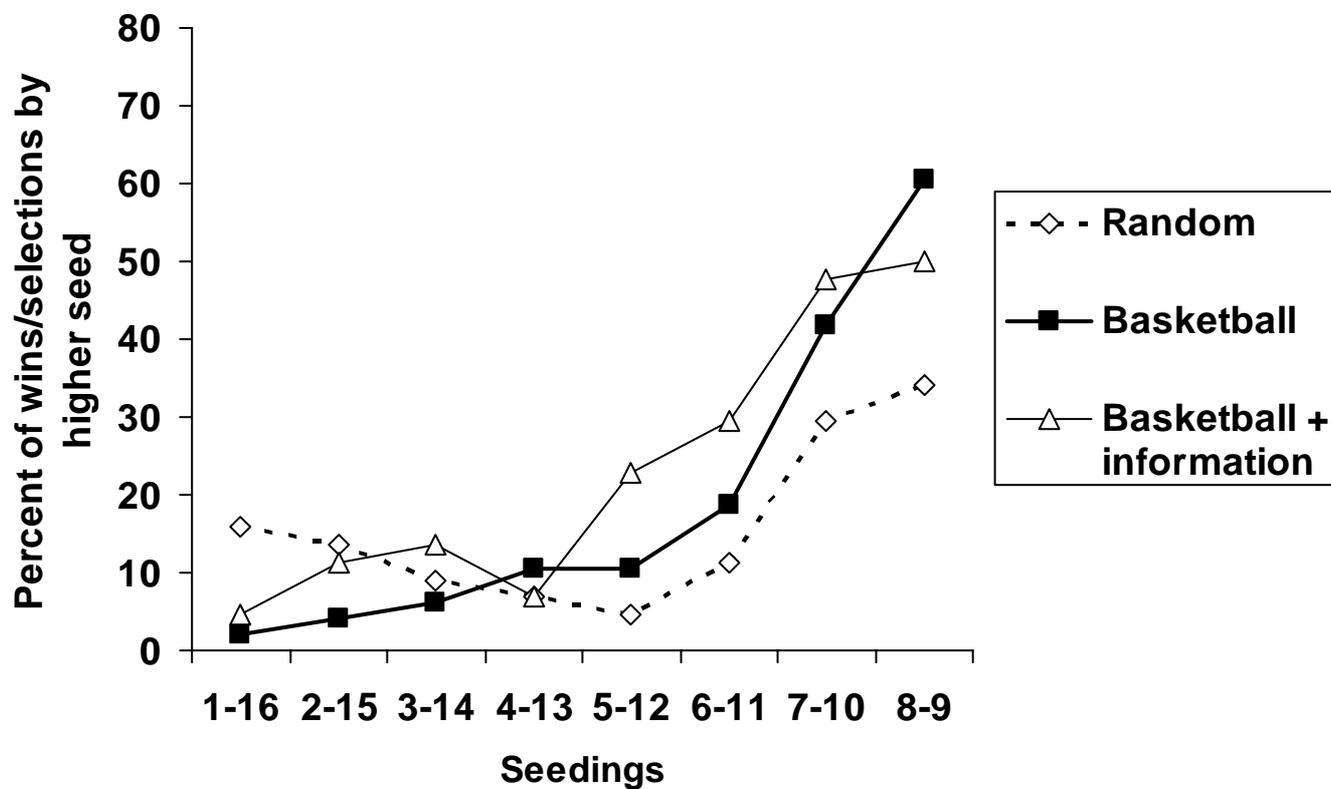
Study 2



Study 2



Study 3



Study 3

