The Role of Experience in the Gambler's Fallacy[†]

Greg Barron* and Stephen Leider

Harvard Business School

and

University of Michigan Ross School of Business

*Address correspondence to: Greg Barron, Harvard Business School, Boston MA, 02163, Email: gbarron@hbs.edu

The Role of Experience in the Gambler's Fallacy

Abstract

Recent papers have demonstrated that the way people acquire information about a decision problem, by experience or by abstract description, can affect their behavior. We examined the role of experience over time in the emergence of the Gambler's Fallacy in binary prediction tasks. Theories of the Gambler's Fallacy and models of binary prediction suggest that recency bias, elicited by experience over time, may play a significant role. An experiment compared a condition where participants sequentially predicted the colored outcomes of a virtual roulette wheel spin with a condition where the wheel's past outcomes were presented all at once. In a third condition outcomes were presented sequentially in an automatic fashion without intervening predictions. Subjects were yoked so that the same history of outcomes was observed in all conditions. The results revealed the Gambler's Fallacy when outcomes were experienced (with or without predictions). However, the Gambler's Fallacy was attenuated when the same outcomes were presented all at once. Observing the Gambler's Fallacy in the third condition suggests that the presentation of information over time is a significant antecedent of the bias. A second experiment demonstrated that, while the bias can emerge with an all-at-once presentation that makes recent outcomes salient (Burns and Corpus, 2004), the bias did not emerge when the presentation did not draw attention to recent outcomes.

The Role of Experience in the Gambler's Fallacy

Consider an expecting mother about to give birth to her 11th child. She currently has 5 boys and 5 girls but the last 4 births have been girls. Understandably she feels strongly that a boy is long "overdue" (i.e. the Gambler's Fallacy). If the mother-to-be describes to her new pediatrician the birth order of her children (MMFMMMFFFF), should we expect the pediatrician to hold the same belief? Imagine now the roulette player at a Vegas casino who has just experienced 5 red outcomes in a row. Even a decision scientist would be hard put not to feel that a black outcome is now likely. But what if the player had just arrived at the table and seen the following history of 10 red (R) and 10 black (B) outcomes displayed above the table (BBRBBBRBBRRRBBBRRRRR). Should she exhibit the Gambler's Fallacy as strongly as the gambler who just witnessed the revelation of those outcomes over time?

In this paper we ask the question: does the *way* we acquire information, by sequential experience or by simultaneous description, play a critical role in the emergence of the Gambler's Fallacy in a binary prediction task? The question is an interesting one since several recent papers on decisions from experience and descriptions suggest that the way people acquire information, by personal experience like the mom-to-be or by description like the pediatrician, can have a significant effect on choice behavior. Binary prediction tasks, like betting on red or black, are a natural decision context to explore for similar effects. By examining this question we hope to both extend the literature on decisions from experience and description, and to deepen our understanding of the underlying cause of the Gambler's Fallacy by identifying conditions that attenuate its emergence.

The Gambler's Fallacy, often attributed to Laplace's essay of 1796[1] and the experimental work of Murray Jarvik (1951), refers to the belief that runs of one binary outcome will be balanced by the opposite outcome. Moreover, the longer the run, the stronger the belief that the opposite outcome is due to appear. Many studies have explored this effect and found it robust in different experimental paradigms such as prediction of binary outcomes, generation of random

---

[1]"I have seen men, ardently desirous of having a son, who could learn only with anxiety of the births of boys in the month when they expected to become fathers. Imagining that the ratio of these births to those of girls ought to be the same at the end of each month, they judged that the boys already born would render more probable the births next of girls."(from A Philosophical Essay on Probabilities by Laplace [1951, p.162])

sequences and identification of sequences as random (e.g. Budescu, 1987; Bar-Hillel & Wagenaar, 1991, for a review, see Lee, 1971, chapter 6).

In this paper we propose that the emergence of the Gambler's Fallacy may be affected by *how* information about past outcomes is presented. Our suggestion is motivated by recent work on decisions from experience and decisions from description. In decisions from description, outcome distributions are described abstractly, for example a choice between $3 with certainty and a lottery providing $32 with probability 0.1 and $0 otherwise. An experience-based choice between the same two options would be based on outcomes incurred in the past from repeated draws from the two distributions. Alternatively, samples may be drawn and merely observed (without incurring any financial gain or loss) after which an individual chooses a single distribution to receive an outcome draw from (see Weber, Blais & Shafir, 2004). In both cases the decision maker has to rely solely on past outcomes to make her decision. Barron & Erev (2003) demonstrate that the deviations from maximization that one observes in choices between lotteries depend critically on how the information was acquired (i.e. through a description or through experience). Notably, while small probabilities are overweighted in decisions from description (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) they tend to be underweighted in decisions form experience (Barron & Erev, 2003; Erev & Barron, 2005; Hertwig, Barron, Weber & Erev, 2004; Fox & Hadar, 2006).

Related research has shown that experience can lead to suboptimal responding in a binary prediction task (Newell & Rakow, 2007). In predicting the binary outcome of a dice throw with four sides of the die in one color and two sides in a second color, participants gave the maximizing prediction more often when the problem was described abstractly, without observing any actual outcomes. Experience, on the other hand, was found to lead to behaviors such as probability matching that are more "representative" of the process generating the outcomes.

Clearly, one would not expect the Gambler's Fallacy to emerge based on an abstract description of a random process. Having been told only that a roulette wheels' red and black outcomes are equally likely; there is no reason to believe one outcome is more likely to appear next. However, it is less clear when predictions are based on a sequence of past outcomes presented all at once. In typical studies of the Gambler's Fallacy subjects experience, and predict, series of binary outcomes one at a time.

We are not aware of published studies where predictions are based solely on fully described sequences of past outcomes (as in the example where a gambler approaches a roulette table and observes the table's history of outcomes). In a somewhat related paradigm Burns and Corpus (2004) observe the Gambler's Fallacy when participants are told to imagine they had observed 100 spins and "the last 4 spins in a row came up red". This stimulus diverges from the full sequences we consider in two ways. First, the last 4 outcomes are portrayed in the aggregate rather than observed as 4 separate outcomes. Second, the other 96 outcomes are not described at all except that they are the outcomes of a random process, and that overall there are an equal number of red and black among the 100 outcomes. A fully *described* sequence of 100 outcomes would constitute a list of all 100 of the actual outcomes. We revisit the importance of this distinction in Experiment 2.

This difference between prediction based on a full sequence and one-at-a-time sequential prediction was the focus of Hogarth and Einhorn's (1992) paper on beliefs and order effects. They undertook a meta-analysis of order effects in studies employing simple tasks with short (2-12 items) series. They find that recency occurred in every study (16/16) where subjects express their beliefs after integrating each piece of evidence in a given sequence Step-by-Step. However, in studies where subjects reported opinions only *after* all the information has been presented, recency was observed much less often (8 out of 27 studies). To organize the pattern of order effects, Hogarth and Einhorn propose a general model of information processing and belief adjustment. For our purposes the key feature is the distinction between Step-by-Step and End-of-Sequence information processing.[2] Hogarth and Einhorn argue that Step-by-Step tasks (i.e. tasks that ask for a response after each piece of evidence) necessitate Step-by-Step thinking, while End-of-Sequence tasks (i.e. a response is only needed after all evidence has been collected) will tend towards End-of-Sequence processing (unless long and/or complex sequences of evidence necessitate Step-by-Step processing due to memory constraints). As the outcomes in our experiment will be simple and the sequence lengths short, we can reasonably expect that

---

[2] Hogarth and Einhorn also distinguish between tasks that call for evaluation (assessing if a hypothesis is true or false) and tasks that call for estimation (constructing some form of a "moving average"). In their model this serves to establish whether the new piece of evidence is compared to a constant reference point (e.g. zero, where an outcome is positive if it supports the hypothesis, and negative if it contradicts) or to the current belief (e.g. last period's posterior for the average). All of our treatments use prediction tasks, which are best classified as estimation tasks. Notably, however, other studies of the Gambler's Fallacy have used identification tasks, which may be considered evaluation tasks in Hogarth's framework.

subjects will use End-of-Sequence thinking when presented with a simultaneous complete description of the previous outcomes. Critically, Hogarth and Einhorn argue that the moving average calculation in Step-by-Step processing will display recency, i.e. increased sensitivity to the last few outcomes, while the holistic End-of-Sequence process will not have recency (and in fact may have primacy, i.e. sensitivity to the initial outcome, due to initial anchoring with a single adjustment for all the evidence).

 We argue, therefore, that applying this adjustment model to either of the leading accounts for the mechanism behind the Gambler's Fallacy (identified by Ayton and Fischer 2004) will naturally imply that the recency bias caused by the sequential presentation is critical for the presence of the Gambler's Fallacy in binary prediction tasks. Estes (1964) presented the first major account of the Gambler's Fallacy, suggesting that subjects bring into the lab a folk intuition that, in general, random outcomes will act like sampling without replacement, based on their experience in the outside world, where such behavior is often the norm. For example, the 100th car in a train portends the caboose with greater likelihood than the third car (Pinker, 1997). When finite populations are sampled without replacement astute observers *should* commit the Gambler's Fallacy. The Gambler's Fallacy bias can then occur when a sequence of the same outcome "uses up" those outcomes from the overall random process. This effect should be more extreme when individuals focus on the smaller subsequence of the most recent results (since smaller sequences are more volatile and thus more likely to deviate from the expected frequencies). For example, imagine (for the sake of argument) that a gambler betting on roulette has a mental model of drawing without replacement from 15 red outcomes and 15 black outcomes. When faced with the sequence BBRBBBRBBRRRBBBRRRRR, if the focus is solely on the 5 most recent outcomes he may believe that Black has a 60% probability of occurring (since there are 10 red and 15 black "left" from the initial 30 outcomes). If, however, he considers all 20 observed outcomes, he correctly believes that red and black are equally likely (5 red and 5 black are "left").

The second proposed model is the representativeness heuristic (Kahneman & Tversky, 1972), where individuals expect that the characteristics of populations are similarly represented at a local level as well. Thus if the likelihood of giving birth to boys and girls is same, the same number of boys and girls is expected in any given small sample of births. Consequently, people expect runs of the same outcome to be less likely than they are. Recency arguably plays a

critical role in the process of determining exactly what is "local" (i.e. the size and serial location of the sample that is expected to be representative of the population). If the recent series of four daughters is particularly salient to the mother, she may feel that the last four outcomes do not reflect the overall equal proportions of males and females; thus she may expect that a boy is much more likely. Alternatively, if the pediatrician considers the entire series, having received it all at once and thus not exhibiting recency, she will likely feel that the outcomes are representative of the population, and therefore would predict that a boy or a girl are equally likely.

This analysis suggests that the way information is encountered will affect the magnitude of the Gambler's Fallacy. Predicting a series of outcomes one-at-a-time makes recent outcomes more salient and is hypothesized to give rise to the Gambler's Fallacy. Predicting an outcome based on the same series presented all at once should draw less attention to recent outcomes and therefore attenuate the fallacy. These predictions are robust to the theoretical choice of the mechanism used to model the Gambler's Fallacy. The following experiment tests this explicitly. In order to examine the necessity of prediction, we also look at the case where outcomes are merely observed as they are revealed in Step-by-Step fashion. Following Hogarth and Einhorn (1992) we hypothesize that the mere presentation of information over time, making recent outcomes more salient, is an important antecedent of the fallacy.

Experiment 1: Three Presentations, Recency and the Gambler's Fallacy

Method

Participants. One hundred and eight volunteers (59 males and 49 females with an average age of 22) served as paid participants in the study. Participants were students (graduate and undergraduate) recruited from several local universities (both liberal arts and professional schools). In addition to the performance-contingent payoff (described below) participants received $15 for participating. The final payoff was approximately $20.

Design, Apparatus and Procedure. Each participant performed a binary prediction task 440 times. Participants were told their task was to predict the outcome of a series of virtual roulette wheel spins whose outcomes could be one of two colors, either red/black or white/blue (see instructions in Appendix A). Participants were shown a window with the past outcomes,

represented as colored balls, of the roulette wheel up to a maximum of 11 (see screenshot in Appendix B).

Participants were randomly allocated to the three experimental conditions, "Sequential," "Simultaneous" and "Autosequential". In the Sequential condition participants predicted sequences of 11 simulated roulette outcomes (only the color) one at a time. After each prediction the next ball was revealed in the history window, a smiley or frowny was presented as additional feedback, and the participant made her next prediction. After the eleventh prediction, feedback was left on screen for 1 second. The history window was then wiped clean and participants began predicting the next sequence of eleven balls. This process was repeated 40 times with participants observing 440 balls (40 x 11) and making 440 predictions (40 x 11) in total.

In the Simultaneous condition, participants predicted only the eleventh outcome after being shown the first ten balls, all at once, in the history window. After making a prediction the next ball was revealed in the history window and a smiley or frowny was presented as additional feedback. One second later, as in the Sequential condition, the history window was wiped clean and participants were given a new series of ten balls. This process was repeated 440 times with participants observing 4400 balls (440 x 10) and making 440 predictions.

The Autosequential condition was the same as the Sequential with one exception. Instead of predicting the first 10 outcomes from the roulette wheel, participants merely watched the balls being revealed one after another at a pace of 1 second per outcome (although it is possible that some subjects may have privately made predictions without prompting). This time interval was picked to approximately equate the total experiment length in each of the conditions[3]. After all 10 balls were revealed participants were asked to predict the next outcome (see instructions in Appendix A). As before, feedback was provided after this prediction, the history was wiped clean and the "new wheel" was presented. Importantly, the same set of 36 random sequences used in the first two conditions was used here.

---

[3] The Autosequential condition was run separately (approximately 2 weeks later) to allow the time interval to be calibrated. Recruitment emails for all sessions were identical and session demographics were similar. Nothing observable to the subjects allowed them to differentiate between the potential sessions. Furthermore, due to the large size of the subject pool (~7000 from multiple large universities) and the method of online recruitment, contamination is not likely, nor is it clear how it would bias results in the current context. Demographics for the condition did not differ significantly from those of the other two conditions. If there were any unobservable differences in this group it's not clear how they might bias our results in the direction we observe.

In all conditions, the sequence of ten balls (recall the eleventh was never revealed) alternated between red/black outcomes and white/blue outcomes to enhance the impression of independent series and roulette wheels as laid out in the instructions.

The series of outcomes observed by participants were prepared in advance by creating a random string of 4400 outcomes using the computer's RND function. Participants in the Sequential condition only saw the first 400 outcomes of their series. Eighteen series were created and then inverted (red became black, white became blue and so on) for a total of 36 series. The inversion was employed to ensure an equal number of both outcomes at the aggregate level. These same 36 series were used in all three experimental conditions. As a result, for each of the first 40 predictions made by a participant in the Simultaneous condition, with 10 balls showing each time, there existed a comparable prediction made by a participant in the Sequential and Autosequential conditions after the same 10 balls were revealed one at a time.

Participants were aware of the expected length of the study (approximately 30 min), so they knew that it included many rounds. To avoid an ''end of task'' effect, they were not informed that the study included exactly 440 trials. Payoffs were contingent on two predictions, randomly selected at the end of the experiment, each of which provided $5 if correct.

Results

In the Sequential condition, subjects made predictions after each outcome, while in the Simultaneous and Autosequential conditions they only made predictions for full sequences of ten outcomes. To make a like-for-like comparison between the conditions, we analyzed predictions for the eleventh outcome, and that were based on the same series of 10 balls in the history window. Because of the yoked random series, these were the first 40 predictions in the Simultaneous condition, every eleventh prediction in the Sequential, and all predictions in the Autosequential condition.

Following Jarvik (1951) and Ayton and Fischer (2004) we calculated the probability, across subjects, that a prediction continues the color of the previous run for runs of length 1 to 4 or more (while all subjects saw runs of 4, 11% never saw runs of 5 or more outcomes). Probabilities smaller than 0.5 represent negative recency, the Gambler's Fallacy. Figure 1 presents the resulting curves for the three conditions. The figure supports this paper's main

hypothesis – participants in the Simultaneous condition appeared to exhibit less negative recency in their predictions than those in the Sequential and Autosequential conditions.

For an initial analysis, we calculated for each subject the mean probability of continuing a run of length 1, 2, 3, and 4 or more. Table 1 reports the mean across subjects, as well as significance levels, from t-tests testing if the observations are different from 0.5. Confirming the impression of Figure 1, we found significant evidence for the Gambler's Fallacy in the Sequential condition, as well as the Autosequential condition for long runs. We did not find evidence for the Gambler's Fallacy in the Simultaneous condition. Following Ayton and Fischer's (2004) analysis, a 3 (information condition) x 4 (run length) analysis of variance (ANOVA) on the subject means confirmed both a significant effect of the information condition [$F(2,420) = 5.04$, MSE = 0.31, $p < 0.01$] on prediction choices, as well as a significant effect of the run length [$F(3, 420) = 2.99$, MSE = 0.18, $p = 0.03$]. There was not a significant interaction effect between condition and run length [$F(6, 420) = 0.55$, MSE = 0.03, $p = 0.77$]. However, we will present below an analysis using a Logit regression, which has greater statistical power, that will find differences between treatments in the significance of the run length variable.

In light of Hogarth and Einhorn's prediction that the all-at-once method of processing information fostered by simultaneous description may lead to primacy effects, we tested whether subject's predictions were correlated with the initial outcome in the sequence. The mean probability (we again first take the average within each subject) of predicting an outcome that matches the first outcome was 0.492 in Simultaneous, 0.484 in Sequential condition and 0.519 in Autosequential. No condition was significantly different from 0.5 (Simultaneous: $t(35) = -0.63$, $p = 0.53$; Sequential: $t(35) = -1.19$, $p = 0.24$; Autosequential: $t(35) = 1.23$, $p = 0.23$). Thus we find no evidence for a primacy effect.[4]

---

[4] We also consider whether other features of the beginning of the sequence, which could affect subjects' perceptions of the random outcome, influence the Gambler's Fallacy. We construct two measures for unbalanced outcomes at the start of the sequence: the length of the longest run in the first 5 observations (an "early run") and the maximum difference between the number of red and black balls (or equivalently blue and white balls) in the first five observations (an "early difference"). For each measure we constructed a dummy variable equal to one if the measure is greater or equal to three. We then performed the same regression as presented in Table 2 with additional interactions between treatment and run length variables and either the "early run" dummy or the "early difference" dummy. While the main Gambler's Fallacy effect pattern remained between treatments, none of the interaction terms with the "early run" dummy were significant, nor were any interactions with the "early difference" dummy.

To more fully analyze our data, we estimated a logit model on choice observations with random effects for each sequence of outcomes. The random effects accounted for the repeated observations on each subject, and the significance levels for each coefficient reflected the fewer numbers of effective independent observations. Table 2 presents the estimated coefficients. We included dummy variables equaling one for predictions in the Sequential and Autosequential conditions (with the constant reflecting the omitted Simultaneous condition). Negative coefficients reflected a propensity for subjects to predict an outcome different from the last outcome, i.e. the Gambler's Fallacy. We also estimated a separate run length coefficient for each treatment: e.g. the variable "Run Length x Sequential" is equal to (run length – 1) for observations in the Sequential condition, and is set to zero for observations in the Simultaneous and Autosequential conditions (with "Run Length x Simultaneous" and "Run Length x Autosequential" defined analogously). Thus we allowed for a separate effect for the length of the run in determining the presence of the Gambler's Fallacy in each treatment. We subtracted one from the run length so that for runs of length one the total estimated Gambler's Fallacy effect could be seen directly from the treatment dummy variables. A negative coefficient on the run length variable means that the Gambler's Fallacy increases for longer runs.

As we expected, the constant term was not significantly different from zero, meaning that the probability that subjects in the Simultaneous condition choose to continue the run was not significantly different from 0.5. Moreover, the coefficient on run length in the Simultaneous condition was also not significantly different from zero, which means that subjects in the Simultaneous condition did not exhibit the Gambler's Fallacy, i.e. did not predict the run to continue with probability less than 0.5, even for long runs. Conversely, the coefficient for the Sequential dummy is significant and negative, meaning that subjects in the Sequential condition had a significant level of the Gambler's Fallacy for all run lengths. Moreover, the coefficient on run length in the Sequential condition was marginally significant and negative, which means that the Gambler's Fallacy was even stronger for long runs. In the Autosequential condition, the dummy variable was not significant, while the run length variable was negative and significant. This should be interpreted to mean that while subjects in the Autosequential condition did not exhibit the Gambler's Fallacy for runs of length one, their probability of continuing the run of outcomes significantly decreased as the length of the run increases. These estimates reflect the results from Figure 1 and Table 1, that the nature of the Gambler's Fallacy is quite different

across the treatments: we find no Gambler's Fallacy for any run length in the Simultaneous condition, the Gambler's Fallacy for both short and long runs in the Sequential condition, and the Gambler's Fallacy for long (but not short) runs in the Autosequential condition.

The results support the hypothesis that sequential presentation of outcomes is an important antecedent of the Gambler's Fallacy. A simultaneous description of past outcomes was found to significantly attenuate the bias. This pattern is consistent with the proposed underlying mechanism, namely that Step-by-Step presentations draw attention to the most recent outcomes in a sequence thereby augmenting the Gambler's Fallacy bias.

To support the proposed mechanism it is useful to consider a *simultaneous* presentation, different to that employed in Experiment 1, which makes recent outcomes salient. If the bias is observed in such a setting it supports the conjecture that recency underlies the observed difference between Simultaneous and Sequential presentations in the current context. Burns and Corpus (2004), summarized above, employed just such a manipulation. While the uniformity of the past 100 roulette outcomes was described, and not experienced sequentially, the description's most salient feature was its emphasis on the most recent outcomes — "the last 4 spins in a row came up red". In this condition, 88% of participants (n=109) predicted black, and thus exhibited the fallacy. While this result supports our interpretation it is not definitive since there are many differences between the paradigm and the Simultaneous condition in the current paper. To address this we replicated the Burns and Corpus (2004) in the experiment below with one minor change. Instead of describing the 100 outcomes in the aggregate with a focus on the most recent outcomes, we present all 100 outcomes. We hypothesize that, lacking the focus on the most recent outcomes, this presentation will not give rise to the bias even though it supplies all of the information available in the Burns and Corpus (2004) presentation.

Experiment 2: Replicating Burns and Corpus (2004) Without a Focus on Recent Outcomes
Method

Participants. Thirty-six volunteers (25 males and 11 females) served as paid participants in the study. Participants were students (graduate and undergraduate) recruited from several local universities (both liberal arts and professional schools). Participants received $10 for participating.

Design, Apparatus and Procedure. Participants were presented with the same scenario from Burns and Corpus (2004) describing the last 100 red or black outcomes of a Las Vegas casino roulette wheel (see full description in Burns and Corpus (2004)). The only change was replacing the sentence "the last 4 spins in a row came up red" with an exhaustive list of the 100 outcomes, represented as the letters R and B. As in the original version, the last 4 outcomes were red (i.e. "R R R R"). Participants were told that outcomes were being presented from left to right with the first and last outcomes clearly labeled. A single question was asked − to predict the next outcome.

Results

The Gambler's Fallacy was not observed, as 18 participants (50%) predicted black and the remainder predicted red. The result supports the hypothesis that recency is an important factor contributing to the Gambler's Fallacy and, presentations which do not *make recent outcomes salient* lead to a significantly attenuated propensity to exhibit the bias.

Discussion

We examined whether experienced information is necessary for the emergence of the Gambler's Fallacy in a binary prediction task. In two experiments participants predicted the colored outcome of a virtual roulette wheel. Information about past outcomes was acquired in one of three ways 1) experienced sequentially, as the feedback from past predictions 2) simply presented sequentially one second at a time, or 3) described simultaneously as a set of past outcomes. The results showed that while the Gambler's Fallacy emerged in the two conditions where information is experienced sequentially over time, there was no tendency towards negative recency when past outcomes are encountered as a description (in the form of a temporally ordered list).

Experiment 2 qualified Burns and Corpus (2004) observation of the Gambler's Fallacy in a simultaneous presentation that drew attention to recent outcomes (88% of subjects exhibited the fallacy). When we used a presentation that contained an exhaustive description (i.e. a list), less likely to make recent outcomes salient, no bias was observed (only 50% predicted a change in color). Taken together, the two experiments suggest that presentations that make recent outcomes salient will augment the fallacy. While experience based presentations are sequential

and naturally draw attention to recent outcomes, descriptive presentations can also give rise to the bias. Here too, as Experiment 2 revealed, recency is an important predictor of the fallacy. Returning to one of the paper's motivating examples, this suggests that while an expecting mother with a past birth order of (MMFMMMFFFF) may tend to believe that a boy is now due, a pediatrician who encounters this information all at once and with no emphasis on the most recent outcomes will not exhibit as strong a tendency.

This paper's main contribution is in identifying the roles of experience and recency in the Gambler's Fallacy. Our results are broadly consistent with Hogarth and Einhorn's (1992) analysis showing that recency emerges robustly when information is revealed Step-by-Step but not when information is acquired all at once. As discussed above, both of the leading mechanisms for the Gambler's Fallacy (Estes 1964; Tversky, 1972) imply that recent observations are a key factor of the bias. However, the results of our first experiment suggest that merely presenting the information sequentially may be sufficient to cause recency effects, even if choices are not required until the end of the sequence. In contrast, Hogarth and Einhorn assume that, if the outcomes are simple and the sequence is short, a task that asks for a choice at the end of the sequence of outcomes will activate a process that does not exhibit recency. Therefore, our results can serve to enrich their model by suggesting a more complete set of criteria to determine when a Step-by-Step process (exhibiting recency) or an End-of-Sequence process (without recency) will be activated. Additionally, we fail to find evidence for a primacy effect in the Simultaneous condition, as predicted by Hogarth and Einhorn's model.

A second account for our results can be derived from an elaboration to Oskarsson, Boven, McClelland, & Hastie's (2008) theoretical work on random and non-random binary sequences and hidden Markov processes. That paper argues that individuals form a mental model of the outcome-generating process that is a Markov process. They argue that the described characteristics of the process (such as whether it is "subjectively random," whether the nonrandom cause is intentional or not, how much control the non-random cause has, and whether the goals are simple or complex) lead individuals to assume different features about the mental Markov process. Depending on the described characteristics, individuals may construct Markov models where the outcomes will be independent over time, or follow trends. They may also conceive of models that shift between states or cycle through states. While their model does not explicitly differentiate within the "subjectively random" category based on the means by which

information is provided (they argue that this category as a whole should elicit the Gambler's Fallacy), their framework is general enough that one could imagine augmenting their model in that direction. Possibly, presenting information sequentially draws more attention to the alternating nature of the series. This may lead to a mental model of a process with a greater emphasis on negative recency (as opposed to independence) and to predictions consistent with this belief. While the first account assumes that experience leads people to overweight recent information, the second assumes that experience leads to exaggerated beliefs about the alternating nature of the sequence. Further work is needed to examine the accounts' relative contributions to the current results.

The current results also contribute to the effort to more sharply define experience. In doing so we seek the minimal set of stimulus characteristics that elicits significantly different behavior than that observed when information is described. We found that passive experience, the observing of outcomes revealed over time, was both necessary and sufficient for the Gambler's Fallacy to emerge. Interestingly, Hertwig et al. (2004) similarly found that the passive sampling of outcomes, without incurring gain or loss, was sufficient in order to observe behavior, such as the underweighting of rare events, which is consistent with past studies of decisions from experience. Taken together these results suggest that, in comparison to behavior based on descriptions, qualitatively different processes are engaged when people merely encounter information sequentially over time. Since this is the way we encounter information in so many contexts (financial, personal, professional, etc.), future research should continue to revisit "classic" behavioral decision theory phenomena using experience-based paradigms. The current results continue to suggest this is a potentially valuable endeavor. Identifying additional contexts where behavior differs from decisions from description will extend the existing literature, and our efforts here, to identify and map the important boundaries between the phenomena of description-based and experience-based decisions.

The conclusions of the current study are limited to the Gambler's Fallacy in *predicted* binary sequences. As noted in the introduction, the Gambler's Fallacy is also observed in generation tasks, where participants are asked to generate a series that looks random, and identification tasks, where series are classified as random or not. Prediction is unique in that generation and identification tasks alert the participant as to the topic of the study (Ayton & Fischer, 2004). Quite possibly, the task of identifying a series as random is an evaluation task (in

the sense of Hogarth and Einhorn, 1992), which is hypothesized not to exhibit recency, rather than an estimation task. If so, we would not expect to find an effect of information presentation in an identification task.

To enable a more sensitive test of when the fallacy occurs we used a paradigm that controls for issues of expected value maximization (in both experiments no one prediction had a higher expected value than another). Consequently, it remains an open question whether or not the findings apply to contexts where the fallacy is actually suboptimal and carries a cost. As noted above, Newell and Rakow (2007) found that, in the die task described above, experience had an initial negative effect on maximization. While serial effects were not reported, our results suggest that the mode of presentation, experience, may have initially augmented the Gambler's Fallacy leading to suboptimal responding. In contrast to our results, passive observation of die rolls (similar to our Autosequential condition) did not lead to greater maximization– a shift was only seen if participants actively predicted outcomes. The interaction between presentation and suboptimality of the fallacy is a question for future research.

More broadly, more work that examines the effect of presentation (simultaneous/descriptive vs. sequential/experiential) and of paradigm (prediction, generation and identification) on behavior is needed (see McDonald & Newell, 2008, for one recent example). One suggestive result is that positive recency (i.e. the hot hand) was observed both in a prediction task with sequential information (NBA free throws) and in an identification task with simultaneously described outcomes (Gilovich, Vallone, & Tversky, 1985).This result is more suggestive than conclusive as it confounds presentation and paradigm (and examines positive recency). However, together with the present study it serves to motivate further study of experience's role in the cognitive biases.

References

Ayton, P., & Fisher, I. (2004). The gambler's fallacy and the hot-handed fallacy: Two faces of subjective randomness. *Memory & Cognition. 32*, 1369-1378.

Bar-Hillel, M., & Wagenaar, W. A. (1991). The perception of randomness. *Advances in Applied Mathematics, 12*, 428-454.

Barron, G., & Erev, I. (2003). Feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*, 215-233.

Budescu, D. V. (1987). A markov model for generation of random binary sequences. *Journal of Experimental Psychology, 13*, 25-39.

Erev, I., & Barron G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, *112,* 912-931.

Estes, W. K. (1964). Probability learning. In A.W. Melton (Ed.), *Categories of human learning* (pp. 88-128). New York: Academic Press.

Fox, C. R., & Hadar, L. (2006). Decisions from experience = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004). *Judgment and Decision Making*, *1,* 159-161.

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

Hertwig, R., Barron, G., Elke, W., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choices. *Psychological Science*, *15*, 534-539.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology, 24,* 1-55.

Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology*, *41*, 291-291.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263-291.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3,* 430-454.

Lee, W. (1971). *Decision theory and human behavior*. New York: John Wiley & Sons, Inc.

Laplace, P. S. (1951). *A philosophical essay on probabilities*. (Translated by F.W. Truscott and F.L. Emory). New York: Dover.

McDonald, F., & Newell, B. (2008). *Effects of alternation rate and prior belief on the interpretation of binary sequences* (University of New South Wales Working Paper).

Newell, B. R., & Rakow, T. (2007). The role of experience in decisions from description. *Psychonomic Bulletin & Review, 14,* 1133-1139.

Oskarsson, A. Van Boven, L., McClelland, G., & Hastie, R. (2008). *What's next? Judging sequences of binary events.* (University of Colorado Working Paper).

Pinker, S. (1997). *How the mind works*. New York: Norton.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *9,* 195-230.

Weber, E. U., Blais, A. R., & Shafir, S. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review*, *111,* 430-445.

Biographical sketches:

Stephen Leider is an Assistant Professor in the Operations and Management Science unit at the University of Michigan, Ross School of Business. Stephen's primary research interests include risk attitudes and social preferences, particularly in organizational and contractual contexts.

Greg Barron is an Assistant Professor in the Negotiation, Organizations, and Markets Unit at the Harvard Business School. Greg's primary research interests are in the field of decision making, focusing on the effects of experience on organizational and consumer behavior.

Figure 1.

Probability of continuing previous run as a function of run length across subjects in Simultaneous, Sequential and Autosequential conditions – 40 yoked trials.
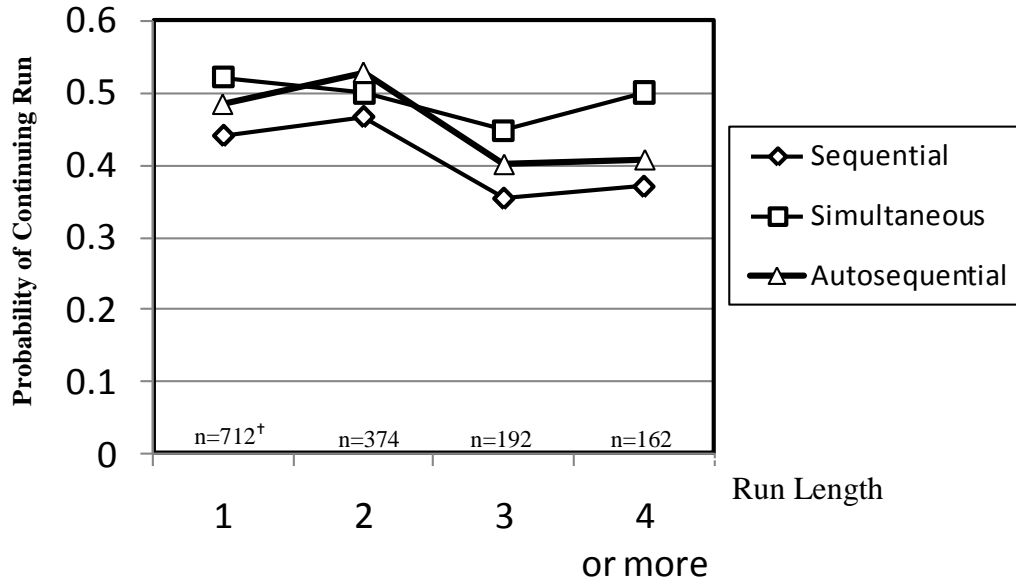
[†] Denotes the number of sequences contributing to each data point for all conditions.

Table 1

Probability of Continuing the Run

Condition	Run Length			
	1	**2**	**3**	**4+**
Sequential	0.438**	0.456	0.369***	0.381**
Simultaneous	0.519	0.506	0.471	0.512
Autosequential	0.484	0.518	0.385**	0.385*

Table 1 reports the mean probability of continuing the run, grouped by subject and run length. Asterisks denote results of a t-test that the probability = 0.5 (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). For Simultaneous: t = 0.85; t = 0.17; t = -0.54 and t = 0.24, respectively. For Sequential: t = -2.67; t = -1.45; t = -3.05 and t = -2.41. For Autosequential: t = -0.72; t = 0.58; t = -2.19 and t = -2.02. All tests are two-tailed and have 35 degrees of freedom.

Table 2

Probability[Choice Continues the Run]	β	Standard Error	Marginal Effect
Dummy for Sequential Condition	-0.253***	0.096	-0.062
Dummy for Autosequential Condition	-0.052	0.095	-0.013
(Run Length) x (Sequential Dummy)	-0.088*	0.048	-0.021
(Run Length) x (Simultaneous Dummy)	-0.031	0.046	-0.008
(Run Length) x (Autosequential Dummy)	-0.109**	0.047	-0.027
Constant	0.044	0.076	
Observations	4248		
Number of series	36		

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2 reports a Logit regression with random effects by outcome series, standard errors and coefficient marginal effects. The dependent variable equals 1 if the subject chose the same outcome as the last outcome in the series, and equals 0 otherwise. The omitted category was the Simultaneous condition. The Run Length variable was defined as (# consecutive outcomes at end of series − 1), thus indicating the effect of longer runs on the Gambler's Fallacy. Wald Test for joint significance: $\chi^2(5) = 25.04$ (p < 0.01). Fraction of variance due to random effects: $\rho = 0.01$.

Appendix A

Participants instructions for the three experimental conditions in Experiments 1 and 2.

Welcome,

In this game your goal is to predict the outcomes of a series of roulette wheels with different pairs of colors.

[SEQUENTIAL: In each round you may choose one of the two colored buttons on the screen. The computer will then randomly generate an outcome that will be one of the two colors. Like a real roulette wheel, the chances for each color being the outcome are equal. A "History" window will display the last 10 outcomes of the game. After every 11 outcomes the game will change over to the next roulette wheel.]

[SIMULTANIOUS: For each round the computer will randomly generate 10 consecutive outcomes that will be displayed from left to right in the "History" window. Like a real roulette wheel, the chances for each color being the outcome are equal. Your task is to then predict the next outcome by choosing one of the two colored buttons on the screen. The outcome will then be randomly generated by the computer. For the next round, the computer will generate 10 new outcomes from the next roulette wheel.]

[AUTOSEQUENTIAL: For each trial, the computer will randomly generate the first 10 outcomes that will be displayed in the "History" window. Like a real roulette wheel, the chance of each of the two colors being the outcome is equal. Your task is to then predict the next outcome by choosing one of the two colored buttons that will appear on the screen. That outcome will then be randomly generated by the computer. The computer will then generate 10 outcomes from the next roulette wheel and again you will make a prediction.]

This process is repeated for a predetermined number of rounds until the experiment is over.

At the end of the experiment, your earnings will depend on only two of the previous rounds, randomly selected by the computer. For each of the two rounds you will receive $5 if you correctly predicted the outcome of that round. Aside from the two rounds, you will also receive $15 for participating.

As in all CLER experiments, deception is not used and all the information above is both true and accurate.

Appendix B

Four screenshots from the experiment.
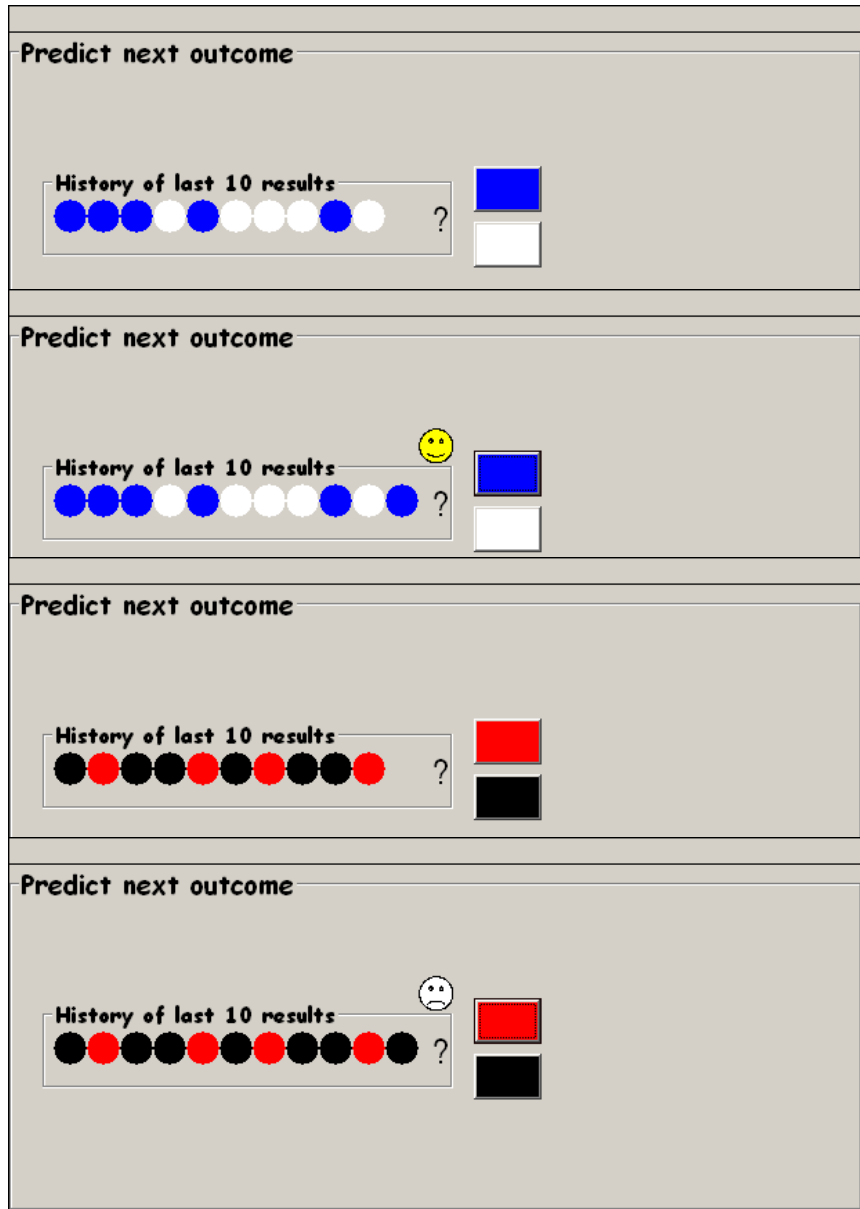
Top screen: A prediction is elicited. Participant chose "Blue", which was correct (second from top). A new set of ten outcomes then appears (third from top). Participant predicts "Red", which was incorrect (bottom screen).